

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 03:59:40 ; Search time 3019 Seconds
(without alignments)
8878.332 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 921

Sequence: 1 atgagtctcaaatgcttat.....acatccaaactggtaactaa 921

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEnbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pt.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	920.2	99.9	921	6	AX180293	Sequence
2	920.2	99.9	1262	6	AX342644	Sequence
3	920.2	99.9	1568	6	AX180299	Sequence
4	908.2	98.6	909	6	AX180295	Sequence
5	771.2	83.7	1671	6	AX360091	Sequence
6	495	53.7	495	6	AX180297	Sequence
7	213.2	23.1	3689	5	XL081291	U81291 Xenopus lae
8	202.8	22.0	164732	9	AC104237	AC104237 Homo sapi
9	202.8	22.0	180707	2	AC012228	AB012228 Homo sapi
10	176.6	19.2	3028	5	AB070367	AB070367 Bufo japo
11	163.2	17.7	2627	6	AX480940	AX480940 Sequence
12	157	17.0	4628	5	XL081290	U81290 Xenopus lae
13	156.2	17.0	169388	2	AC129620	AC129620 Rattus no
14	114.6	12.4	2913	6	AX360086	Sequence
15	89.4	9.7	169388	2	AC129620	Sequence
16	83.6	9.1	2025	10	AF356627	AF356627 Mus muscu
17	80.4	8.7	2231	10	BC019485	BC019485 Mus muscu
18	76.2	8.3	2409	6	AX360098	Sequence
19	74.6	8.1	2672	6	AX207905	Sequence
20	74.6	8.1	3104	6	AX207903	Sequence
21	72.2	7.8	1850	5	AB038496	AB038496 Xenopus l
22	71	7.7	891	4	BOVERCS	L19663 Bos taurus
23	71	7.7	2581	6	I64576	I64576 Sequence 1
24	71	7.7	3860	4	BTU09859	U09859 Bos taurus
25	69.4	7.5	3586	4	PIGENPR	D30799 Sus scrofa
26	68.8	7.5	708	6	AX207901	Sequence
27	68.8	7.5	853	3	AF130842	AF130842 Rhyzopert
28	68.8	7.5	3534	3	CIN431686	AF131686 Clona Int
29	68.4	7.4	2594	3	AY119618	AY119618 Drosophil
30	68.4	7.4	2752	10	NMU300738	AJ300738 Mus muscu
31	68.4	7.4	3749	3	DROSTUBBLE	L11451 Drosophila
32	65.8	7.1	942	3	SSU79521	U79521 Scolopendra
33	65.6	7.1	3696	6	AX207959	Sequence
34	65.6	7.1	3696	9	HSU09860	U09860 Human enter
35	65.2	7.1	2331	4	AF395821	AF395821 Oryctolag
36	64.6	7.0	2407	3	AY052044	AY052044 Drosophil
37	63.2	6.9	1916	9	AF045649	AF045649 Homo sapi
38	63.2	6.9	2087	9	HUMFXI	M1142 Human facto
39	62.4	6.8	829	3	AF487426	AF487426 Aedes aeg
40	62.2	6.8	3183	10	BC029645	BC029645 Mus muscu
41	60.8	6.6	891	3	AF508783	AF508783 Aedes aeg
42	60.2	6.5	1303	10	AB032200	AB032200 Mus muscu
43	60	6.5	2230	4	AB022425	AB022425 Sus scrof
44	59.8	6.5	3652	3	DMU18130	U18130 Drosophila
45	59	6.4	813	4	P1GFIXA	M26235 Pig factor

ALIGNMENTS

RESULT 1	AX180293	AX180293	921 bp	DNA	linear	PAT 06-AUG-2001
LOCUS	Sequence 1	from Patent WO0146407.				
DEFINITION	AX180293					
ACCESSION	AX180293					
VERSION	AX180293.1	GI:15132262				
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
REFERENCE	1 (bases 1 to 921)					
AUTHORS	Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and Sands,A.T.					
TITLE	Polynucleotides encoding human protease homologs					

JOURNAL Patent: WO 0146407-A 1 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers
source 1..921

BASE COUNT 242 a 195 c 240 g 242 t 2 others
ORIGIN

Query Match 99.9%; Score 920.2; DB 6; Length 921;
Best Local Similarity 100.0%; Pred. No. 8.6e-277;
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCAAAATGCTTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
Db 1 ATGAGTCTCAAAATGCTTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
Qy 61 TTGGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTCCTGGGACAGT 120
Db 61 TTGGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTCCTGGGACAGT 120
Qy 121 CTGTTTAAGGTACAGCTTGGAAATTTTAAACATTTTTCAGTGCATCTCTGGAGGAGC 180
Db 121 CTGTTTAAGGTACAGCTTGGAAATTTTAAACATTTTTCAGTGCATCTCTGGAGGAGC 180
Qy 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGACATATT 240
Db 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGACATATT 240
Qy 241 TGTGGAGGAAGCAATCTCAACACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300
Db 241 TGTGGAGGAAGCAATCTCAACACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300
Qy 301 AGAACATTTGCTACTTTTGAATGTTTACCTGGAGAGTATGACITTAAGCCAGACAG 360
Db 301 AGAACATTTGCTACTTTTGAATGTTTACCTGGAGAGTATGACITTAAGCCAGACAG 360
Qy 361 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCATACATACATCCACATTTCTCCACCAAG 420
Db 361 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCATACATACATCCACATTTCTCCACCAAG 420
Qy 421 AAACCAATGACATATGATATGCGCTTTTGAAGTGGCTGGAGCTTCCCAATTTGGCCAC 480
Db 421 AAACCAATGACATATGATATGCGCTTTTGAAGTGGCTGGAGCTTCCCAATTTGGCCAC 480
Qy 481 TTTGTGGGCCCCATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTTATTGT 540
Db 481 TTTGTGGGCCCCATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTTATTGT 540
Qy 541 ACAACTGCAGGCTGGGCGCTTTAACTGAAGTGGCTCTCTCACAAGTCTTTGCAGGAA 600
Db 541 ACAACTGCAGGCTGGGCGCTTTAACTGAAGTGGCTCTCTCACAAGTCTTTGCAGGAA 600
Qy 601 GTGAATCTGCATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 660
Db 601 GTGAATCTGCATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTGTGTAAACACTAAAGAGG 660
Qy 661 CCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTTCCTGATGGAGGAGAGCAGCATGT 720
Db 661 CCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTTCCTGATGGAGGAGAGCAGCATGT 720
Qy 721 CAGGAGATTTCAGAGGTTTCACATCATGTGCGGAAATGAAGAGGGGCTTGACTCTGGCT 780
Db 721 CAGGAGATTTCAGAGGTTTCACATCATGTGCGGAAATGAAGAGGGGCTTGACTCTGGCT 780
Qy 781 GGTGTGACTTCTCGGGGTTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAAGT 840
Db 781 GGTGTGACTTCTCGGGGTTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAAGT 840
Qy 841 GATCAAGGATCCCTGGATCTTCACAGACATTTAGTAAAGTCTTTCTCGATCCAGGAA 900
Db 841 GATCAAGGATCCCTGGATCTTCACAGACATTTAGTAAAGTCTTTCTCGATCCAGGAA 900

Qy 901 CACATCCAAACTGGTAACTAA 921
Db 901 CACATCCAAACTGGTAACTAA 921
RESULT 2
LOCUS AX342644 1262 bp DNA linear PAT 12-JAN-2002
DEFINITION Sequence 41 from Patent WO0198468.
ACCESSION AX342644
VERSION AX342644.1 GI:18152041
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Yue, H., Elliott, V.S., Gandhi, A.R., Lal, P., Au-Young, J.,
Tribouley, C.M., Deleage, A.M., Baughn, M.R., Nguyen, D.B., Lee, E.A.,
Hafalia, A., Khan, F.A., Wallia, N.K., Yao, M.G., Lu, D.A., Patterson, C.,
Tang, Y.T., Walsh, R.T., Azimzai, Y., Ramkumar, J., Xu, Y. and Reddy, R.
JOURNAL Patent: WO 0198468-A 41 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..1262
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 433459CB1"
BASE COUNT 354 a 265 c 306 g 337 t
ORIGIN
Query Match 99.9%; Score 920.2; DB 6; Length 1262;
Best Local Similarity 99.8%; Pred. No. 8.9e-277;
Matches 919; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGTCTCAAAATGCTTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
Db 165 ATGAGTCTCAAAATGCTTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 224
Qy 61 TTTGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTCCTGGGACAGT 120
Db 225 TTTGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTCCTGGGACAGT 284
Qy 121 CTGTTTAAGGTACAGCTTGGAAATTTTAAACATTTTTCAGTGCATCTCTGGAGGAGC 180
Db 285 CTGTTTAAGGTACAGCTTGGAAATTTTAAACATTTTTCAGTGCATCTCTGGAGGAGC 344
Qy 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGACATATT 240
Db 345 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGACATATT 404
Qy 241 TGTGGAGGAGCAATCTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300
Db 405 TGTGGAGGAGCAATCTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 464
Qy 301 AGAACATTTGCTCTACTTTGAATGTTTACCTGGAGAGTATGACTTAAAGCCAGACAG 360
Db 465 AGAACATTTGCTCTACTTTGAATGTTTACCTGGAGAGTATGACTTAAAGCCAGACAG 524
Qy 361 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCATCATACATCCATTTCTCCACCAAG 420
Db 525 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCATCATACATCCATTTCTCCACCAAG 584
Qy 421 AAACCAATGACATATGATTTGCGCTTTTGAAGTGGCTGGAGGCTTCCAAATTTGGCCAC 480
Db 585 AAACCAATGACATATGATTTGCGCTTTTGAAGTGGCTGGAGGCTTCCAAATTTGGCCAC 644
Qy 481 TTTGTGGGCCCCATATGCTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTATTGT 540
Db 645 TTTGTGGGCCCCATATGCTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTATTGT 704
Qy 541 ACAACTGCAGGCTGGGCGCTTTAACTGAAGTGGGCTCTCTCACAAGTCTTTGCAGGAA 600
Db 541 ACAACTGCAGGCTGGGCGCTTTAACTGAAGTGGGCTCTCTCACAAGTCTTTGCAGGAA 600

Db 705 ACAATGACGGCTGGGGCCGCTTAACCTGAAGGTGGGTCCTCTCACAAAGCTTTGCAGAA 764

Qy 601 GTGAATCTGCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTTAACACTAAAGAGG 660

Db 765 GTGAATCTGCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTTAACACTAAAGAGG 824

Qy 661 CCATCAGTGGGAGACCTTTCTTTTGACAGTGTTCCTGATGGAGGAGAGACGCATGT 720

Db 825 CCATCAGTGGGAGACCTTTCTTTGCACAGTGTTCCTGATGGAGGAGAGACGCATGT 884

Qy 721 CAGGAGATTACAGAGTTTACATCATGTGCGCGAATAAGAAAGGGCCTGGACTCTGGCT 780

Db 885 CAGGAGATTACAGAGTTTACATCATGTGCGCGAATAAGAAAGGGCCTGGACTCTGGCT 944

Qy 781 GGTGTGACTTCCCTGGGCTTTGGGCTGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAGT 840

Db 945 GGTGTGACTTCCCTGGGCTTTGGGCTGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAGT 1004

Qy 841 GATCAAGGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCCCTGGATCCACGAA 900

Db 1005 GATCAAGGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCCCTGGATCCACGAA 1064

Qy 901 CACATCCAAACTGGTAACTAA 921

Db 1065 CACATCCAAACTGGTAACTAA 1085

RESULT 3

AX180299

LOCUS AX180299 1568 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 7 from Patent WO0146407.

ACCESSION AX180299

VERSION AX180299.1 GI:15132265

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1568)

TITLE Walke, D.W., Turner, C.A., Abuin, A., Friedrich, G., Zambrowicz, B. and Sands, A.T.

JOURNAL Polynucleotides encoding human protease homologs

PATENT: WO 0146407-A 7 28-JUN-2001;

Lexicon Genetics Incorporated (US)

FEATURES

Location/Qualifiers

1..1568

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 452 a 338 c 353 g 423 t 2 others

ORIGIN

Query Match 99.9%; Score 920.2; DB 6; Length 1568;

Best Local Similarity 100.0%; Pred. No. 9.1e-277;

Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAAGCTGATTTTACTACTAGGAATAGTCTTT 60

Db 245 ATGAGTCTCAAAATGCTTATAAGCAGGAAAGCTGATTTTACTACTAGGAATAGTCTTT 304

Qy 61 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCAAGCTCCCAAGTCTTGGGCAGAGT 120

Db 305 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCAAGCTCCCAAGTCTTGGGCAGAGT 364

Qy 121 CTGGTTAAGTACAGCTTGAATTTTAAACATTTTCACTTCAGTCGCATTTTGGAGGAGC 180

Db 365 CTGGTTAAGTACAGCTTGAATTTTAAACATTTTCACTTCAGTCGCATTTTGGAGGAGC 424

Qy 181 CAAGTGGAGAGGTTTCTATCCCTGGCAGGTATCTCTGAAACAAAGGCAAGACATATT 240

Db 425 CAAGTGGAGAGGTTTCTATCCCTGGCAGGTATCTCTGAAACAAAGGCAAGACATATT 484

Qy 241 TGTGGAGGAAGCATCGCTCTCACCACAGTGGGTGATCAGCGCGCTCAGCTCATTTGCAAAAC 300

Db 485 TGTGGAGGAAGCATCGCTCTCACCACAGTGGGTGATCAGCGCGCTCAGCTCATTTGCAAAAC 544

Qy 301 AGAAACATTTGCTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360

Db 545 AGAAACATTTGCTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 604

Qy 361 CCAGGAGAGCAAACTCTCAGTATTGAAACCTGTTCATCATACATCCACATTTCTCCACCAAG 420

Db 605 CCAGGAGAGCAAACTCTCAGTATTGAAACCTGTTCATCATACATCCACATTTCTCCACCAAG 664

Qy 421 AAACCAATGGAGCTATGATATTGCCCCCTTTTGAAGATGGCTGGAGGCCCTTCCAATTTGGCCAC 480

Db 665 AAACCAATGGAGCTATGATATTGCCCCCTTTTGAAGATGGCTGGAGGCCCTTCCAATTTGGCCAC 724

Qy 481 TTTGTGGGGCCCCATATGCTCTCCAGAGCTGCGGAGCAATTTTCAGAGCTGCTGTTTTATTGT 540

Db 725 TTTGTGGGGCCCCATATGCTCTCCAGAGCTGCGGAGCAATTTTCAGAGCTGCTGTTTTATTGT 784

Qy 541 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGGCTCCTCTCACAAAGTCTTTCGAGGAA 600

Db 785 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGGCTCCTCTCACAAAGTCTTTCGAGGAA 844

Qy 601 GTGAATCTGCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTTAACACTAAAGAGG 660

Db 845 GTGAATCTGCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTTAACACTAAAGAGG 904

Qy 661 CCCATCAGTGGGAGACCTTTCTTTGCACAGTGTTCCTCTGATGGAGGAGAGACGCATGT 720

Db 905 CCCATCAGTGGGAGACCTTTCTTTGCACAGTGTTCCTCTGATGGAGGAGAGACGCATGT 964

Qy 721 CAGGAGATTACAGAGTTCACATCATGTGCGGAAATAAGAAAGGGCCTGAGCTCTGGCT 780

Db 965 CAGGAGATTACAGAGTTCACATCATGTGCGGAAATAAGAAAGGGCCTGAGCTCTGGCT 1024

Qy 781 GGTGTGACTTCCCTGGGCTTTGGGCTGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAGT 840

Db 1025 GGTGTGACTTCCCTGGGCTTTGGGCTGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAGT 1084

Qy 841 GATCAAGGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCCCTGGATCCACGAA 900

Db 1085 GATCAAGGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCCCTGGATCCACGAA 1144

Qy 901 CACATCCAAACTGGTAACTAA 921

Db 1145 CACATCCAAACTGGTAACTAA 1165

RESULT 4

AX180295

LOCUS AX180295 909 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 3 from Patent WO0146407.

ACCESSION AX180295

VERSION AX180295.1 GI:15132263

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 909)

TITLE Walke, D.W., Turner, C.A., Abuin, A., Friedrich, G., Zambrowicz, B. and Sands, A.T.

JOURNAL Polynucleotides encoding human protease homologs

PATENT: WO 0146407-A 3 28-JUN-2001;

Lexicon Genetics Incorporated (US)

FEATURES

Location/Qualifiers

1..909

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 237 a 193 c 238 g 239 t 2 others

ORIGIN

Query Match 98.6%; Score 908.2; DB 6; Length 909;

Best Local Similarity 100.0%; Pred. No. 4.9e-273;


```

Db 122 AATTGAACCTACCTTTCTCGAATTGTTGGAGGGAGAGAATCAAGAAAGGTCACACATCCAT 181
QY 206 GGCAGGTATCTCGAACAAGGAGAGAGCATATTTGTGGAGGAGCATCGTCTCACCAC 265
Db 182 GGACAGTATCATTTGAACGGAATGGAAGCATTTTGTGGAGGAACCCCTGTGAGCCATT 241
QY 266 AGTGGGTGATCAGCGCGCTCACTGCAATTCGAACAGAAACATTTGTCTACTTTTGAATG 325
Db 242 GTCATGTACTACAGTGCACCATGTTTACTCGACAGGATGTGAATTTATATATGAGAG 301
QY 326 TTACTGCTGGAGAGTATGACTTAAGCAGACAGACCCAGGAGACAACTCTCCTCATTTG 385
Db 302 TTTACATTTGGAGATATGATCAAAATCTTAAAGAAAGAACTTCAACAGATGTTCCAGGGTTA 361
QY 386 AAACGTGTCATACATCCACATTTCTCCACCAGAAACCAAGGACTATGATATGCCCC 445
Db 362 TCGAAATCTTCAACACCCCACTTTTAAACAGAGTCAGCCCAATGATGATGTTGCTG 421
QY 446 TTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGGCCCATTTTGTGGGCCCATATGTCTTCCAG 505
Db 422 TACTGTTGCTGGATGTTCCGTAACTTTTGATGAGAATATTCACACGAGCATGCTGCGCAA 481
QY 506 AGCTGCGGGAGCAATTTAGGCTGGTTTATTTGTACAACTGCAGGCTGGGGCGGCTTAA 565
Db 482 ACCAGATGATGTTTGAACCGGGTGACCTCTGTGTTACTTTGGGTGGGGTCATTTAA 541
QY 566 CTGAAGGTGGCGCTCTCTCAACAGCTCTGCAGGAAGTGAATCTGCCTATTTTGCACCTGGG 625
Db 542 CAGAGATGGAATCTCCCTGTTGTTTACAGAGGCTCTACTTACCCATTTGTGGACCTCT 601
QY 626 AAGAGTGTGGCAGCTCTTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTT 685
Db 602 CATGCTGCTGCTATGTCATGCTAGCTCTGAAGGACCTGTAAGGAACCGTGTGTCTTCTATATAGTGT 661
QY 686 GCACAGGTTTCTGATGGAGGAGACGCATCTCAGGGAGATTCAGGAGGTTCACTCA 745
Db 662 GTGCAGGTTTCCAGAAAGGGAAGGAGCCCTGCCAGGGGACTCAGGAGGCCGTTGC 721
QY 746 TGTGCCGAATAAGAAAGGGCGCTGGACTCTGGCTGTGTGTGACTTCTCGGGGTTTGGGCT 805
Db 722 TGTGCCAGCGCGTCATGGAAGTTGGTTTCTGTCATGCTGTGACCTCTCGGGGATGGGT 781
QY 806 GTGTCAGGCTGGAGAAACAATG-----TGAGGAAAGATGATCAAGGATCCCTTGGGA 859
Db 782 GTGACCTTCATGGAATAATGTTGTTTGTGCTCATATAGGAAGGATCTCCAGGCA 841
QY 860 TCTTCACAGACATAGTAAAGTGTCTTCCCGGATCCACGAACACATCCAACTGGT 915
Db 842 TTTTCACAGACATCCAGAAATGCTTGGCTGGTGGCTCCAGTCCGAGTTGAATACAGCT 897

RESULT 8
AC104237 164732 bp DNA linear PRI 30-APR-2002
LOCUS Homo sapiens chromosome 11, clone RP11-35J10, complete sequence.
DEFINITION AC104237
ACCESSION AC104237.2 GI:20128277
VERSION HTG.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164732)
REFERENCE Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S.,
Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C.,
LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164732)
REFERENCE
AUTHORS
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepeil,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164732)
REFERENCE
AUTHORS
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepeil,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 10, 2002 this sequence version replaced gi:17386405.
All repeats were identified using RepeatMasker.

```

TITLE
 JOURNAL
 COMMENT

```
repeat_region 25237. 25470
/rpt_family="L2"
repeat_region complement(25810. 25919)
/rpt_family="MIR"
repeat_region 28709. 29154
/rpt_family="LTR16A1"
repeat_region 29252. 29284
/rpt_family="(CAAAA)n"
repeat_region 29612. 30220
/rpt_family="L1"
```

```
repeat_region
30415..30485
/rpt_family="L1"
30485..30788
repeat_region
```

```
repeat_region
/rpt_family="L1"
30486..30788
/rpt_family="AluX"
repeat_region
complement(30809..31550)
/rpt_family="L1p"
repeat_region
31551..32627
/rpt_family="L1PA3"
repeat_region
complement(32836..33451)
/rpt_family="L1p"
```

```
repeat_region complement(34836, .33431)
/rpt_family="L1P"
repeat_region complement(33489, .33965)
/rpt_family="L1"
repeat_region 33977, .34152
/rpt_family="L1"
repeat_region complement(34157, .34342)
/rpt_family="MER46A"
repeat_region 34343, .35088
/rpt_family="L1Mda"
repeat_region 35089, .35303
/rpt_family="L2"
repeat_region 35968, .36617
```

```

repeat_region      /rpt_family="L1PA13"
unsure             /note="single clone coverage"
very Match         22.0%; Score 202.8; DB 9; Length 164732;
Local Similarity   92.6%; pred.No. 9.9e-52;
Matches 213; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

	Local Similarity	92.6%;	pred. No. 9.9e-32;	Mismatches 213;	Conservative	0;	Mismatches 17;	Indels	0;	Gaps	0;
692	GTTTTCTTCGATGGAGGGAGACGCCA	NTCAGGAGATTTCAGGAGGTTCACTCATGTGCC	751								
11264	GTGATGCTTAATAGAGCGTTTTTCCCT	CTGCAGGAGATTTCAGGAGGTTCACTCATGTGCC	51323								
752	GGAATAAAGAAAAGGGGCCCTGGACT	CTGGCTGGTGTCACATCTCTGGGGGTTGGGCTGTGGTC	811								
1324	GGAATAAAGAAAAGGGGCCCTGGACT	CTGGCTGGTGTCACATCTCTGGGGGTTGGGCTGTGGTC	51383								

51324 GGAATTAAGAAAGGCGCTTGGACTTGGCTGGTGTGACTTCTTGGGTGTGGCGTGTGGTC 51383

812 GAGCGTGGAGAAACAATGTGAGGAAAAGTATCAAGGATCCCTCGGGATCTTCACAGACA 871

51384 GAGCGTGGAGAAACAATGTGAGGAAAAGTATCAAGGATCCCTCGGGATCTTCACAGACA 51443

872 TTATGTAAGTGCCTTCTCGGATCCACGAAACACATCCCAAACCTGGTAACATA 921

51444 TTATGTAAGTGCCTTCCCTCGGATCCACGAAACACATCCCAAACCTGGTAACATA 51493

TTAGTAAAGTGCTCCCTGGATCCACGACATCCCAACTCGGTAACATA 51493

AC012228 180707 bp DNA linear HTG 17-AUG-2002
Homo sapiens chromosome 11 clone RP11-439A13 map 11, *** SEQUENCING
IN PROGRESS ***, 2 ordered pieces.
AC012228

AC012228.8 GI:22123492
 HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 180707)
 Birren,B., Nussbaum,C. and Lander,E.

1 (bases 1 to 180707)
Birren, B., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone RP11-439A13
Unpublished
2 (bases 1 to 180707)

```

Every Match      22.0%; Score 202.8; DB 9; Length 164732;
1st Local Similarity 92.6%; Pred. No. 9.9e-52;
Matches 213: Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

692 GTTTTCCGTGATGGAGGGAGACGCATGTCAGGGAGATTCAAGAGGTTCACTCATGTGCC 751
|| | ||| ||| | |||||||||
11264 GTGATGCTAATAAGAGCGTTTTTCCTCTGAGGGAGATTCAAGAGGTTCACTCATGTGCC 51323

752 GGAATAAGAAAGGGGCTGGACACTCTGGCTGGTGTGACTTCCTGGGGTTTGGGCTGTGCTC 811

51324 GGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCCTGGGGTTGGGCTGTGGTC 51383

812 GAGGCTGGAGAAACAATGTGAGGAAAAAGTGATCAAGGATCCCGCTGGGATCTCTTCAACAGACA 871

51384 GAGGCTGGAGAAACAATGTGAGGAAAAGTGATCAAGGATCCCTGGGATCTTTCACAGACA 51443

872 TTAGTAAAGTGCCTTCCCTGGATCCACGAACACATCCAAACTGGTAACATA 921
|||||
1144 TTAGTAAAGTGCCTTCCCTGGATCCACGAACACATCCAAACTGGTAACATA 51493

2228/c

5 AC012228 180707 bp DNA linear HTG 17-AUG-2002
 NITION Homo sapiens chromosome 11 clone RP11-439A13 map 11, *** SEQUENCING
 IN PROGRESS ***, 2 ordered pieces.

SESSION AC012228

ION AC012228.8 GI:22123492

HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.

human.

ANISM Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

Eukaryota, Metazoa, Chordata, Vertebrata, Eucereotes comit,
Mammalia: Eutheria: primates: Catarrhini: Hominiidae: Homo.

REFERENCE 1 (bases 1 to 180707)

THORS Birren, B., Nusbaum, C. and Lander, E.

TLTLE Homo sapiens chromosome 11, clone RP11-439A13

Unpublished

REFERENCE 2 (bases 1 to 180707)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Doyle,M., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Gage,D., Ferrera,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,J., Peterson,J., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE

Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,I., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21362200.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

REFERENCE

AUTHORS

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2182
Center clone name: 439_A_13

NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 71362: contig of 71362 bp in length
* 71363 71462: gap of 100 bp
* 71463 180707: contig of 109245 bp in length.
Location/Qualifiers
1. .180707
/organism="Homo sapiens"
/db_xref="taxon.9606"
/chromosome="11"
/map="11"
/clone="RP11-439A13"
/clone_11b="RP11-439A13"
47741 a 39327 c 40651 g 52888 t 100 others

FEATURES

source

FEATURES

source

1. .3028
/organism="Bufo japonicus"
/db_xref="taxon.8387"
/sex="female"
/tissue_type="oviductal pars recta"
85. .3009
/function="cleaves egg envelope glycoproteins"
/note="oviduct serine protease"
/codon_start=1
/product="oviductin"
/protein_id="BAB63372.1"
/db_xref="GI:15277254"
CDS

CDS

ORIGIN

Query Match 22.0%; Score 202.8; DB 2; Length 180707;
Best Local Similarity 92.6%; Pred. No. 1e-51;
Matches 213; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 692 GTTTCCTCTGAGGAGACGCATGTCTAGGGAGATTTCAGGAGGTTCTACTCATGTGCC 751
Db 165369 GTGATGCTAATAGAGCTTTTCCTCTGCAGGAGATTTCAGGAGGTTCTACTCATGTGCC 165310
QY 752 GGAATAAGAAAGGGCTGGACTCTGGCTGGTGTGACTCTCTGGGTTTGGGCTGTGGTC 811
Db 165309 GGAATAAGAAAGGGCTGGACTCTGGCTGGTGTGACTCTCTGGGTTTGGGCTGTGGTC 165250
QY 812 GAGCTCGAGAAACAATGTGAGGAAAAGTGATCAAGGATCCCTGGGATCTTCACAGCA 871
Db 165249 GAGCTCGAGAAACAATGTGAGGAAAAGTGATCAAGGATCCCTGGGATCTTCACAGCA 165190
QY 872 TTAGTAAGTGTCTTCTCGATCCACGAACACATCCAAACTGGTAACTAA 921
Db 165189 TTAGTAAGTGTCTTCTCGATCCACGAACACATCCAAACTGGTAACTAA 165140

RESULT 10

AB070367

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FCGFPALDPVLVSVMQIITFTSDYSANYLGFRAVIVISFVLPSSPVKPEKGNQPRKNO
DAMQHFEGCGVSLPFRFLYHNLKAEAMPNSPWHVSINFGNKHVCNGAILSKTF
VWTSANCAVDREPERPSIGLIVAGLHDLESSINTOKRPEYVIVHPDYNRLSKDIDVAL
IHVORPQYNSYVOPICGLPDGHSRLSPSKLCVSGWDLNVELSKLOOLEVPVLMDDV
CKKYIDTDMFCVAGIAEDNASCLAQSGAPLVCSAPGYAIFGIVSRVGCNET
PKAGYSSVFLFIPWIMETILSVAGIIDTSEPHPLFPDPKPSQOKALLPDSPPSS
SODIYVCKDVLSPQELKVASGDGPGEGRCQLIFOAPEGHFILLFKQLSHEH
YSLIIEGASSNKTFFKAQLMEEKIPTIMKSAGAVITILEASTAQDSALHLWLSYSEHN
QN"

misc_feature 232..990
/note="encodes protease domain"
misc_feature 1018..1359
/note="encodes CUB domain"
misc_feature 1387..1725
/note="encodes CUB domain"
BASE COUNT 800 a 696 c 727 g 805 t
ORIGIN
Query Match 19.2%; Score 176.6; DB 5; Length 3028;
Best Local Similarity 53.1%; Pred. No. 1e-43;
Matches 402; Conservative 0; Mismatches 349; Indels 6; Gaps 1;

Qy 158 TCAGTCGCATCTTGGAGAGCCCAAGTGGAGAGGGTTCCTATTCCTGGCAGGTATCTC 217
Db 224 TCTCTCGTATTGTTGGAGGGACATCTGCAGTGAAAGGGGAAGCCCATGGATGGTCTG 283
Qy 218 TGAACAAAGGAGAGAGCATATTTCTGGAGGAAGCATCTCTCACACAGTGGGTGATCA 277
Db 284 TAAAAGAGATGGGAAGCACATCTCTGGAGGAACCATTTATAGCGATAAATACGTGCTTA 343
Qy 278 CGGGGCTCACTGCTATTCGAACAGAAACATTGTCTTACTTTTGAATCTTACTGCTGG 337
Db 344 CGCTGCGCATTCGCTCTTAGAAAAAACTTTGAATTTCAAGTGAGCGCTTCGATGGAG 403
Qy 338 AGTATGACTTAAAGCAGACAGACCCAGGAGAGCAAACTCTCACATTTGAAACTGTCA 397
Db 404 ATCAGATTTCGCTGTATATGAAAGGAGTGAACAGCGTTTCGCCATAAAATCTGTTTCA 463
Qy 398 TACATCCACATTTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTTGAAGATGG 457
Db 464 AGCACCCCACTTCAACCGAGCAGGCCCTTCAACTATGACTCGCCATACTGGAGCTGG 523
Qy 458 CTGAGCGCTTCCAATTTGGCCACTTTGTGGGGCCCATATGTCTTCCAGAGCTGCGGGAGC 517
Db 524 TGGAAAGCATCACATTTGATAAAGATATCCAAACCGCATGCTTACCGAGCCCGATGATG 583
Qy 518 AATTGAGCGCTTTTATTTGTACAACTGCAGGCTGGGCGGCTTAAGTCAAGTGGCG 577
Db 584 TATTCCAAAGAACTCTGTATGGCACTTGTGTGGGAGCTGTGCGAGGAAAATGAA 643
Qy 578 TCCTCTCACAGTCTTGAGGAAGTGAATCTGCCTTATTTGACCTTGGGAAGAGTGTGTGG 637
Db 644 GGTACCTAGCAGCTGCAAAAAGTTGTTTACCTCTCATTTGAGTACAGAAGTGTCTTGA 703
Qy 638 CAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAACACCTTCTTTTGCACAGTTTTC 697
Db 704 GCATCATGGAGACTGTGGATCGGCTTTTGGCATTTTGAGACTGTGGTTTGTGCAAGATTTC 763
Qy 698 CTGATGGAGGAGAGACCATGTTCAGGAGATTCAGGAGGTTTCACTCATGTGCGGGAATA 757
Db 764 CAGAAAGGGGTAAAGACCGGTGTACAGGTGACTTGGAGGTCTTCTCTGCGCCAGAGA 823
Qy 758 AGAAAGGGGCTGAGCTGTGGTGTGACTTCTCTGGGGTTTGGGCTGTGGTGTGAGGCT 817
Db 824 GCCAAGAGGTGGGTTTGGTGTGTAACTCTCTGGGGATAGGATGTGCACGAAAT 883
Qy 818 GGAGAAACAATGTGAGGAAAAGTG-----ATCAAGGATCCCTGGGATCTTCAAGACA 871
Db 884 GGGCAGACAATATCTGTGACCCAGTGGAAAGCAAAAGTTTACCAGGTGTTTTTTCACGGATA 943
Qy 872 TTAGTAAGTCTTCTTCCTGGATCCAGNACACATCCA 908
Db 944 TTCAGAGGCTTCTCAACTGGCTCAGTGAACAACTCAA 980

RESULT 11
AX480940
LOCUS AX480940
DEFINITION Sequence 36 from Patent WO0246383.
ACCESSION AX480940
VERSION AX480940.1 GI:22217579
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
Yue, H., Azimzai, Y., Kallick, D. A., Baughn, M. R., Griffin, J. A.,
Swarnakar, A., Lal, P. G., Walla, N. K., Hafalia, A. J., Gandhi, A. R.,
Au-Young, J., Elliott, V. S., Ramkumar, J., Thangavelu, K., Lu, Y.,
Warren, B. A., Lee, E. A., Tribouley, C. M., Arvizu, C.,
Deleane, A. M., Yao, M. G., Khan, F. A. and Sanjanwala, M. M.
Protein modification and maintenance molecules
Patent: WO 0246383-A 36 13-JUN-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1..2627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7484157CBI"
BASE COUNT 739 a 561 c 636 g 691 t
ORIGIN

Query Match 17.7%; Score 163.2; DB 6; Length 2627;
Best Local Similarity 51.7%; Pred. No. 1.6e-39;
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;
Qy 148 TTTAATATTTTCAGTCGCATCTTGGAGGAAGCCCAAGTGGAGAGGGTTCCTATCCCTGG 207
Db 156 TCTAGATTTCTCTAGAAATTAGTAGTTGGAGAAATTCACACAGTCTGATCCATCGG 215
Qy 208 CAGGTATCTCTGAACAAAGGAGCAGAGCATATTTGTGGAGGAGCATGCTCTCACACAG 267
Db 216 CAGGTCTCCCTAAATCAGATGAGCACCCTCTCTGTGGAGGAGCTTGATTCAGAAAG 275
Qy 268 TGGTGATCAGCGGCGCTCACTGTCATTTGCAACAGCA-----AACATTGTGTCTACTTTG 321
Db 276 CGGTTGTTTACAGCAGCACACTGCTGGACAGCCTCAGTGAGACAGCTGAAGAATATA 335
Qy 322 AATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACT 381
Db 336 ACTGTGACTTCTGGGGAGTACAGCCTCTTTTTCAGAAAGGATAAGCAAGACAGAAATATCCT 395
Qy 382 ATTGAAATCTCATCATACATCCACATTTCTCCCAAGAAACCAATGGACTATGATATT 441
Db 396 GTCTCAAAATATTATCCCATCTCTGAATACACACCCGTGAATATATGAGTCTGATATT 455
Qy 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTTGTGGGCCCCATATGTCTT 501
Db 456 GCAGTGTCTATCTAAAAACACAAAGTCAAGTTTGGAAATGCTGTTACAGCAATCTGTCTT 515
Qy 502 CCAGAGCTGCGGGAGCAATTTGAGCCTGTTTATTTTGTACAACTGCGAGCTGGGGCGCC 561
Db 516 CTTGACAGCGATGATAAAGTTTGAACCGAATTTCTCTTATCCAGTGGATGGGGCAAG 575
Qy 562 TTAATGAAGTGGCGTCTCTCAACAGTCTTTCAGAGGAAGTGAATCTGCCTATTTTGTACC 621
Db 576 ATTTCCAAAGACATCAGAATATTCAATCTCTACAAAGAAATGGAACTTCCCATCATGAT 635
Qy 622 TGGGAAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACTTTT 681
Db 636 GACAGAGCGTGTACTACTGTCTCAAGAGCATGAACCTCCCTCCCTCGGAGGAGCATG 695
Qy 682 CTTTGCACAGGTTTCTCTGATGGAGGAGAGCAGCATGTCAGGGAGATTCAGAGGTTTCA 741
Db 696 CTGTGTCTGGCTTCCCTGATTTGGGGAATGGAGCCTGCCAGGGGGGACTCTGGAGGACCA 755

QY	742	CTCATGTGCGCGGAATAGAAGGGCCCTGCACACTCTGGCTGGTGTGACTCTCTCCCTGGGGTTTG	801
Db	756	CTGCTGTTGTAGAAGAGGTGGTGAATCTGATGATCTTCTGCTGGGATAACTCTCTGGGTAGCT	815
QY	802	GGCTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAAGTGCATCAAGGATCCCTGGGATC	861
Db	816	GGTTGCTGAGGCTTCAGTTCCTCCGTTAAGAACAAACCATGTGAAGGATCCTTGGCATT	875
QY	862	TTTACAGACATTAGTAAAGTGCTTTCTCTGGATCCAGCAACACATCCAAACTGGT	915
Db	876	TTCTCCAAAGTGTGAGTTGATGGATTTTATCACTCAAAACCTGTTACAGGT	929
RESULT 12			
XLU81290			
LOCUS	XLU81290	4628 bp	mRNA linear VRT 11-OCT-2000
DEFINITION	Xenopus laevis polyprotein, serine proteases and ovoidinase regions, mRNA, complete cds.		
ACCESSION	U81290 U44951		
KEYWORDS	U81290.1 GI:2981640		
SOURCE	Xenopus laevis.		
ORGANISM	Xenopus laevis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopus; Xenopodinae; Xenopus.		
AUTHORS	1 (bases 1 to 4628)		
TITLE	Lindsay,J.L., Yang,J.C. and Hedrick,J.L.		
JOURNAL	Ovoidinase, a Xenopus laevis egg extracellular protease, is translated as part of an unusual polyprotease		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11253-11258 (1999)		
PUBMED	99432219		
REFERENCE	2 (bases 1 to 4628)		
AUTHORS	Yang,J.C., Lindsay,L.L. and Hedrick,J.L.		
TITLE	cDNA Cloning of Ovoidinase, a Chymotrypsin-like Protease Released From Xenopus laevis Eggs at Fertilization		
JOURNAL	Unpublished		
REFERENCE	3 (bases 3737 to 4608)		
AUTHORS	Yang,J.C. and Hedrick,J.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-DEC-1995) Molecular and Cellular Biology, University of California at Davis, Davis, CA 95616, USA		
REFERENCE	4 (bases 3736 to 4628)		
AUTHORS	Yang,J.C., Lindsay,L.L. and Hedrick,J.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-DEC-1996) Molecular & Cellular Biology, University of California, Davis, CA 95616, USA		
REFERENCE	5 (bases 1 to 4628)		
AUTHORS	Yang,J.C., Lindsay,L.L. and Hedrick,J.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAR-1998) Molecular & Cellular Biology, University of California, Davis, CA 95616, USA		
REMARK	Nucleotide and amino acid sequence updated by submitter		
COMMENT	On or before Oct 11, 2000 this sequence version replaced		
gi:1216404, gi:1754711.			
FEATURES	Location/Qualifiers		
source	1..4628		
	/organism="Xenopus laevis"		
	/db_xref="taxon:8355"		
	11..4585		
	/codon_start=1		
	/product="polyprotein"		
	/protein_id="AAC24717.1"		
	/db_xref="GI:2981641"		
	/translation="MNVSITLLIWAACISAKISNLEPOQESPPIQETAEKCGTRPE		
	IGDPLETTSRTVGGDAAGVQPTVSLKLNHERHICGSIVRKDMVYTAACHVCPV		
	TEKVSMTFTVIVEGDVQWDSQEOSIPVSHIEPHPNYRGDNGMDIALVFLSPIL		
	FSGVQVPLCPQVEKLEAGTLCVSGWHLRENGDLSPLQVEKLVVVDNGTCHAVL		
	EPIGHVPLDDTMLCAGPEGGMDACGDSGGPFCRRNRSGVFLACGVSVMGCGRSV		
	GAKQITRSQSGSPAFISRYFLPKPLGSGSGKRTITGKNGTVRYPLSGNYS		
	INSVCRWMLQAKATIEIRFLQDIEDHATCTFDYLSFTVNEKMKIRKVCGSTIPSP		
	L		
CDS			

5

```

IYRSNKVTVTFSDCTFTGGRPEIQFLAIPTKAASACGSAKILKKKGMIYSNYPDPPI
PRUKTSWIEAPENHIVKLEDFNVEIGHGCIIDAVEYDIOAEKQLIARLCUGTII
PLPISSEPMNLIIRFKTDNEMSPGKVFKEFSLPVDVDTPTTISPMWILPRAIIA
LDCVGMAPMTPKWMLPRIIVGEEASPNSPWQVQIFLRTFHEGAILSPWMLTAAH
CIRAAEPSYWTIVAGDHRLMNESTEQIRNIKTIRHDNINSETYDNDIALLYLEEP
DLYNDFVRPCPEPEPVLTPASVCVVTWGNTAEDGQAPALGQLOQLPLDLSIICNTS
YISGELDHLMLCAGFPSSKEKDAQDGGSLVCQNEKEQFSIYGLVSWGEGCGRVK
PGYVTVLFTWQIQDQDLQOENALLNSKSVQREGRKAAGCCSDVLUVELVLIA
SPGYGAGGAGLRCSWVIEAAPFSLIKLTKDLSIBESDCLNEFLATLYNIDKVK
TRGSLHNNKTVSSAQRVHFQSGDGGSNRNOGFFLOYWIVSVQVPLAFTGLSQ
GPRYGSSCEDVILTSEGVETSPNLYGNYPDDLHCQWRIIAPGVKIVRVDLLELKEK
DVSQCOODPLVYNGIGESKDLGVGCDEIRLSIMKSEGESEILIFITSNSEVSGKFS
KYSFWDQAGSKOLENAAVGLVCAJDLQIPSEMGSALESPPFPPLVSNEMDCQW
ITYSKYKGLQEVLYNLEAFSPNCSDWYINIDGPNNSHLLASLCGHQDVLHNS
GTFTVLFHSDSVSDWGFKIQGVAPPEPSDTLRSULDSHESDSEDEPSETKKQK
QCGNTVDPMLYLAHCEVKNLRNKKGKESGRVGGQQAAPRSWPLVSIQNSKKRHY
CGGLITNKWILTAAHCEVKNLRHVGVGHTDLTEVQNEHAKVYNHSHVELYMPGSS
PRNDLLELDTPLLNSVAVIICLPDDVTDTQAECLVAGWGVTDVGMGLPTKLQ
WAGPVISTKCKDQYVWSVDVTDNNICAKAGASCMSGDGSLICKREDRYIYLVGV
QKSGKCDKVRAPSYVTLTSAFMDWISQHMDT"
179. .934
/product="serine protease"
mat_peptide 1760. .2461
/product="serine protease"
mat_peptide 3893. .4582
/product="ovochoymase"
/function="role in the blocking of polyspermy"
/note="egg chymotrypsin-like protease; located in the
perivitelline space of the unactivated egg and is released
at fertilization"
941. .1273
/note="encodes a CUB domain"
misc_feature 1295. .1633
/note="encodes a CUB domain"
misc_feature 2533. .2881
/note="encodes a CUB domain"
misc_feature 2948. .3289
/note="encodes a CUB domain"
misc_feature 3341. .3700
/note="encodes a CUB domain"
BASE COUNT 1257 a 1025 c 1166 g 1180 t
ORIGIN
Query Match 17.0%; Score 157; DB 5; Length 4628;
Best Local Similarity 53.0%; Pred. No. 1.5e-37;
Matches 362; Conservative 0; Mismatches 315; Indels 6; Gaps 1;
Qy 163 CGCATTCCTTGAGGAAGCGAAGTGAGGAAGGGTTCCTTATCCCTGGCAGGATCTCTGAAA 222
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 176 CGCATTTGGGAGGAGGAGATGCTGCTGTTGGAGGACACCATGGCGGTGCTCGTNAAG 235
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 223 CAAAGGCAGAGCATATTTGTGGAGGAAGCATCGTCTCACACAGTGGGTGATCACGGCG 282
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 236 CTGAATGAGAGGCATATCTGTGGGGCGGAGCATTTGTGGCGAAGGACATPGTGGTAACTGCG 295
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 283 GCTCACTGCGAT-----TGCACACAGAAACATCTGTGCTACTTTGATGTACTGCTGGA 336
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 296 GCCCATCTGTGTCTATCTGTGACGTAAAGTGAAGTGAGCCACATGACTGTGATTTGTTGGC 355
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 337 GAGTATGACTTTAAGCCACAGACAGCCACGAGGACAAACTCTCTCACTATTTGAAACTGTCTATC 396
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 356 GAATATGACCAGCAAGTTATGGACAGTCAAGACGACAGACATCCCGTATCCCATATAGAG 415
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 397 ATACATCCACATTTCTCCACAGAAACCAATGAGCATATGATATGCGCTTTTGAAGATG 456
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 416 CCGCACCACAAATACAGAGGGGATTTGGAAGATGGGCTATGACATTTGCTTTGCTTTCTG 475
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 457 GCTGAGCGCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGTCTTCCAGAGCTGCGGGAG 516
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 476 TCTAAGCAATCATCTTTTGGCTCCAGAGTGCAACCCCAATTTGGCCTTTCCGAGGTGGGAGAG 535
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 517 CAATTTGAGGCTGTTTTATTATTACAACTGCAGGCTGGGGCGCGCTTAACGTGAAGGTGGC 576
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```


Db 536 AAGATTGAAGCGGACCTGTGTGTTCCAGCGCTGGGAGGCTTGAAGAAAATGGG 595
QY 577 GTCCCTCTCACAAGTCTTGCGAGGAAGTGAATCTGCCTATTTCACCTGGGAAAGAGTGTGTG 636
Db 596 GACTTGTGCCAGTCTTACGAGGAAGTAAACATGCCAGTCTGTGATTAATGGCACCTGCCCAT 655
QY 637 GCAGCTCTCTTAACACTAAAGAGGCGCATCATGCTGGAAGACCTTTCTTTTGCACAGGTTTTT 696
Db 656 GCTGTGTGGAGCCCATAGGCGCATCCGTTATTAGATGACACAATGTTATGTGCTGGATTTC 715
QY 697 CTGTATGGAGGAGAGACCCATGTCAGGAGATTCAGGAGGTTCACTATGTCGCGGAAT 756
Db 716 CCAGAGGAGGATGATGATCCCTGCCAGGCGATTCTGGAGGACCCCTTTGTATGTCAGAGCG 775
QY 757 AAGAAAGGGCGCTGCTGCTGTGTGACTTCTCTGGGCTTTGGCGTGTGTCGAGGC 816
Db 776 CGCTCTGGGGTGTGTTCTTGTGCTGGATGTGTATCTCTGGGCGCTTGGCTGTGGGAGAAC 835
QY 817 TGGAGAAACATGTTGAGGAAAG 839
Db 836 TGGGGTGCACAAACAGATCATACG 858

RESULT 13
AC129620
LOCUS
DEFINITION
AC129620 169388 bp DNA linear HTG 31-JUL-2002
Rattus norvegicus clone CH230-388116, *** SEQUENCING IN PROGRESS
***, 56 unordered pieces.
AC129620
AC129620.1 GI:22024303
HTG: HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 169388)
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbala,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brleva,M., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,F.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsone,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwar,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwu,S., Ogih,M., Okwuonu,G.,
Orgunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Prinus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Unpublished
2 (bases 1 to 169388)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAHL
Center clone name: CH230-388116
----- Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 136807 bases at least Q40
Consensus quality: 141381 bases at least Q30
Consensus quality: 145526 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1002: contig of 1002 bp in length
* 1003 1102: gap of unknown length
* 1103 2132: contig of 1030 bp in length
* 2133 2232: gap of unknown length
* 2233 3358: contig of 1126 bp in length
* 3359 3459: gap of unknown length
* 5024 5123: contig of 1565 bp in length
* 5124 6663: contig of 1540 bp in length
* 6664 6764: gap of unknown length
* 6764 8065: contig of 1302 bp in length
* 8066 8165: gap of unknown length
* 8166 9312: contig of 1147 bp in length
* 9313 9412: gap of unknown length
* 9413 11980: contig of 2568 bp in length
* 11981 12080: gap of unknown length
* 12081 13407: contig of 1327 bp in length
* 13408 13507: gap of unknown length
* 13508 15457: contig of 1950 bp in length
* 15458 15557: gap of unknown length
* 15558 16567: contig of 1010 bp in length
* 16567 16667: gap of unknown length
* 16668 17971: contig of 1304 bp in length
* 17972 18071: gap of unknown length
* 18072 19740: contig of 1669 bp in length
* 19741 19840: gap of unknown length
* 19841 21942: contig of 2102 bp in length
* 21943 22042: gap of unknown length
* 22043 23410: contig of 1368 bp in length
* 23411 23510: gap of unknown length
* 23511 25146: contig of 1636 bp in length
* 25147 25246: gap of unknown length
* 25247 26357: contig of 1111 bp in length
* 26358 26458: gap of unknown length
* 26459 28550: contig of 2093 bp in length

* 28551 28650: gap of unknown length
* 28651 29929: contig of 1279 bp in length
* 29930 30029: gap of unknown length
* 30030 31427: contig of 1398 bp in length
* 31428 31527: gap of unknown length
* 31528 33329: contig of 1802 bp in length
* 33330 33429: gap of unknown length
* 33430 34854: contig of 1425 bp in length
* 34855 34954: gap of unknown length
* 34955 36498: contig of 1544 bp in length
* 36499 36598: gap of unknown length
* 36599 39161: contig of 2563 bp in length
* 39162 39261: gap of unknown length
* 39262 40905: contig of 1644 bp in length
* 40906 41005: gap of unknown length
* 41006 44218: contig of 3213 bp in length
* 44219 46887: gap of unknown length
* 44319 46887: contig of 2569 bp in length
* 46888 46987: gap of unknown length
* 46988 49280: contig of 2293 bp in length
* 49281 49380: gap of unknown length
* 49381 51242: contig of 1862 bp in length
* 51243 51342: gap of unknown length
* 51343 53327: contig of 1985 bp in length
* 53328 53427: gap of unknown length
* 53428 55718: contig of 2291 bp in length
* 55719 55818: gap of unknown length
* 55819 59313: contig of 3495 bp in length
* 59314 59413: gap of unknown length
* 59414 61947: contig of 2534 bp in length
* 61948 62047: gap of unknown length
* 62048 64296: contig of 2249 bp in length
* 64297 64396: gap of unknown length
* 64397 66012: contig of 1616 bp in length
* 66013 66112: gap of unknown length
* 66113 68206: contig of 2094 bp in length
* 68207 68306: gap of unknown length
* 68307 71222: contig of 2916 bp in length
* 71223 71322: gap of unknown length
* 71323 74509: contig of 3187 bp in length
* 74510 74609: gap of unknown length
* 74610 77284: contig of 2675 bp in length
* 77285 77384: gap of unknown length
* 77385 80659: contig of 3275 bp in length
* 80660 80759: gap of unknown length
* 80760 84060: contig of 3301 bp in length
* 84061 84160: gap of unknown length
* 84161 86617: contig of 2457 bp in length
* 86618 86717: gap of unknown length
* 86718 89029: contig of 2312 bp in length
* 89030 89129: gap of unknown length
* 89130 92826: contig of 3697 bp in length
* 92827 92926: gap of unknown length
* 92927 97425: contig of 4499 bp in length
* 97426 97525: gap of unknown length
* 97526 100539: contig of 3014 bp in length
* 100540 100639: gap of unknown length
* 100640 106363: contig of 5724 bp in length
* 106364 106463: gap of unknown length
* 106464 111809: contig of 5346 bp in length
* 111810 111909: gap of unknown length
* 111910 115701: contig of 3792 bp in length
* 115702 115801: gap of unknown length
* 115802 121844: contig of 6043 bp in length
* 121845 121945: gap of unknown length
* 121945 130099: contig of 8155 bp in length
* 130100 130199: gap of unknown length
* 130200 137395: contig of 7196 bp in length
* 137396 137495: gap of unknown length
* 137496 143948: contig of 6453 bp in length
* 143949 144048: gap of unknown length
* 144049 151651: contig of 7603 bp in length
* 151652 151751: gap of unknown length

* 151752 160340: contig of 8589 bp in length
* 160341 160440: gap of unknown length
* 160441 169388: contig of 8948 bp in length.
FEATURES
source
 I..169388
 /organism="Rattus norvegicus"
Query Match 17.0%; Score 156.2; DB 2; Length 169388;
Best Local Similarity 86.1%; Pred. No. 4e-37;
Matches 173; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 721 CAGGGAGATTACAGGAGTTTCACTCATGTGCCGGAATAAGAAAGGGGCTGTGACTTGGCT 780
 |||||
DB 98817 CAGGGAGATTACAGGAGGTCGCTTATCTGTCAAGATAGGAAGGGGCTGTGACTTACT 98876
QY 781 GGTGTGACTTCTCGGGGCTTTGGGCTGTGTCGAGGCTGGAGAACAAATGTGAGGAAAAGT 840
 |||||
DB 98877 GGAGTGACTTCTTTGGGCTGTGGCTGTGGTCGAGCTGGAGAAACAATGCAAGGAAGAAA 98936
QY 841 GATCAAGGATCCCTCGGATCTTCACAGACATTAAGTAAAGTGTCTTCTGGATCCACGAA 900
 |||||
DB 98937 GAACAAGGATCTCTCGGATCTTCACAGATCTTAGGAGAGTGTCTTCTGGATCCATGAA 98996
QY 901 CACATCCAAACTGGTAACTAA 921
 |||||
DB 98997 CACGTCCAGACTGGTAATGAA 99017
RESULT 14
AX360086 2913 bp DNA linear PAT 13-FEB-2002
LOCUS
DEFINITION
ACCESSION AX360086
VERSION AX360086.1 GI:18675712
KEYWORDS
SOURCE human.
ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS
 Florman,G., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
 Charyczak,G.
TITLE
 Novel proteases
JOURNAL
 Patent: WO 0200860-A 42 03-JAN-2002;
 Sugen, Inc. (US)
FEATURES
source
 I..2913
 Location/Qualifiers
BASE COUNT 819 a 590 c 698 g 806 t
ORIGIN
Query Match 12.4%; Score 114.6; DB 6; Length 2913;
Best Local Similarity 52.4%; Pred. No. 2.7e-24;
Matches 252; Conservative 0; Mismatches 229; Indels 0; Gaps 0;
QY 435 TGATATTGCCCTTTTGAAGATGGCTGGAGCTTCCAATTTGGCCATTTTGTGGGCCCAT 494
 |||||
DB 9 TGATATTGCACTGCTGTATCTAANAACAAGTCAAGTTTGGAAATGCTGTTCAGCCAT 68
QY 495 ATGTCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTGTATTTTGTACAACATGCAGGCTG 554
 |||||
DB 69 CTGTCTTCTTGACAGCGATGATAAAGTTGAACACAGGAATTTCTTGTCTTATCCAGTGGATG 128
QY 555 GGGCGGCTTAACGAAGTGGCGCTCTCACAGTCTTCAGGAAGTGAATCTGCCTAT 614
 |||||
DB 129 GGGCAAGATTTCCAAACATCAGAAATATTCAAATGTCTTACAGAAATGGAACATTCCTAT 188
QY 615 TTTGACCTGGGAAGAGTGTGTGGCAGCTCTCTTTAACTAAAGAGGCCCATCAGTGGGAA 674
 |||||
DB 189 CATGGATGACAGAGCGTGAATACTGTGCTCAAGAGCATGAACCTCCCTCCCTCCCTGGGAG 248
QY 675 GACCTTTCTTTGCACAGGTTTCTCTGATGGAGGAGACGCATGTGAGGAGATTCAGG 734

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 04:03:10 ; Search time 275 seconds
(without alignments)
7542.147 Million cell updates/sec

Title: US-09-735-713a-1

Perfect score: 921

Sequence: 1 atgagtctcaaatgcttat.....acatccaactggtactaa 921

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	920.2	99.9	921	AAD09328	Human protease hom
2	920.2	99.9	1262	ABK12903	Human protease PRT
3	920.2	99.9	1568	AAD09331	Human protease hom
4	908.2	98.6	909	AAD09328	Human protease hom
5	771.2	83.2	1671	ABK31790	DNA encoding novel
6	546.2	59.3	637	AAH98934	Toad EST-derived c
7	495	53.7	495	AAD09330	Human protease hom
8	260	28.2	324	AAC05816	Human secreted pro
9	163.2	17.7	2742	AAS88433	DNA encoding novel

c	10	163.2	17.7	3309	23	AAS87759	DNA encoding novel
	11	163.2	17.7	4729	23	AAS72057	DNA encoding novel
	12	163.2	17.7	4729	23	AAS74405	DNA encoding novel
	13	163.2	17.7	4729	23	AAS78775	DNA encoding novel
	14	163.2	17.7	4729	23	AAS84074	DNA encoding novel
	15	161.6	17.5	3096	24	ABQ86158	Novel human gene.
	16	114.6	12.4	2913	24	ABK31785	DNA encoding novel
	17	76.2	8.3	1327	20	AZ34034	Human EST DNA35597
	18	76.2	8.3	1327	21	AAC78495	Human EST DNA35597
	19	76.2	8.3	1327	21	AAC58237	Human EST (express
	20	76.2	8.3	2409	24	ABK31797	DNA encoding novel
	21	76.2	8.3	3143	20	AAC34033	Human PRO618 nucle
	22	76.2	8.3	3143	21	AAC78494	Human PRO618 (UNQ3
	23	76.2	8.3	3143	21	AAC58236	Human PRO618 nucle
	24	74.6	8.1	2672	22	AAD13117	Human membrane-typ
	25	74.6	8.1	3104	22	AAD13116	Human membrane-typ
	26	71	7.7	708	21	AAK99151	708nt DNA sequence
	27	71	7.7	708	24	ABK51438	DNA encoding cattl
	28	71	7.7	2581	15	AAQ70104	Bovine enterokinas
	29	68.8	7.5	708	22	AAD13115	Human membrane-typ
	30	68.4	7.4	3736	23	ABL28971	Drosophila melanog
	31	65.6	7.1	3696	22	AAD13165	Human enterokinas
	32	64.6	7.0	3196	23	ABL26453	Drosophila melanog
	33	63.2	6.9	2087	24	ABK83923	Human cDNA differe
	34	54.6	5.9	973	24	AAL45562	Lumbricus rubellus
	35	54.4	5.9	936	22	AAF75659	Human ztrypl degen
	36	54	5.9	796	24	ABK30337	Human G-protein-co
	37	53.2	5.8	734	24	ABK16738	Epithelin-like ser
	38	53.2	5.8	927	24	ABK16740	Epithelin-like ser
	39	53.2	5.8	2267	24	ABN95751	Gene #2249 used to
	40	53.2	5.8	2267	24	ABL67214	Thyroid cancer rel
	41	53.2	5.8	3387	24	ABK31781	DNA encoding novel
	42	53.2	5.8	3711	24	ABK86141	DNA encoding huma
	43	53	5.8	973	22	AAD03445	Human plasminogen
	44	53	5.8	1476	22	AAS64180	Human prostate cDN
	45	53	5.8	1476	22	AAH93944	PI000C open readin

ALIGNMENTS

RESULT 1
AAD09328
ID AAD09328 standard; cDNA; 921 BP.
XX AAD09328;
XX
DT 10-SEP-2001 (first entry)
DE Human protease homologue cDNA #1.
KW Human; protease homologue; novel human protein; NHP; therapy;
KW pharmacogenomic application; physiological disorder; ss.
XX Homo sapiens.
XX
FH Key
FT CDS
FT
FT Location/Qualifiers
FT 1..921
FT /*tag= a
FT /product= "Human protease homologue #1"
FT 67..69
FT /*tag= b
FT /note= "Encodes Arg"
FT 68
FT misc_feature
FT /*tag= c
FT /note= "This degenerate base represents a
FT polymorphic site"
FT 82..84
FT /*tag= d
FT /note= "Encodes Ala"
FT 82
FT misc_feature
FT /*tag= e
FT /note= "This degenerate base represents a

PI Azimzal Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
DR WPI: 2002-090437/12.
DR P-PSDB: AAU74760.
XX
XX Twenty one human proteases (referred to as PRS-1 to PRS-21), useful
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
PT proliferative (e.g. cancer) disorders -
XX
XX Claim 5; Page 175; 177pp; English.
XX
XX The present invention relates to twenty one new human proteases,
CC referred to as PRS-1 to PRS-21. The PRS polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
CC myocardial infarction, autoimmune/inflammatory e.g. acquired
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
CC endometriosis disorders. Numerous other examples of each disorder are
CC given in the specification. The present nucleic acid sequence encodes
CC the human protease PRS-20 protein of the invention.
XX
XX Query Match 99.9%; Score 920.2; DB 24; Length 1262;
XX Best Local Similarity 99.8%; Pred. No. 6.7e-285;
XX Matches 919; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 ATGAGTCTCAAAATGCTTATAGCAGGAGCAAGCTGATTTTACTACTAGGATAGTCTTT 60
Db 165 ATGAGTCTCAAAATGCTTATAGCAGGAGCAAGCTGATTTTACTACTAGGATAGTCTTT 224
Qy 61 TTTGAACRAGGTAAATCTGCAACCTTTTCGCTCCCAAGCTCCAGTGTGGGAGAGT 120
Db 225 TTTGAACRAGGTAAATCTGCAACCTTTTCGCTCCCAAGCTCCAGTGTGGGAGAGT 284
Qy 121 CTGGTTAAGGTACAGCTTGGAAATATTTTAAACATTTTCACTGCGCATTTTGGAGGAAGC 180
Db 285 CTGGTTAAGGTACAGCTTGGAAATATTTTAAACATTTTCACTGCGCATTTTGGAGGAAGC 344
Qy 181 CAAGTGGAGAGAGGTCTCTATCCCTGGCAGGTATCTCTGAACAAAGCGCAGACATATT 240
Db 345 CAAGTGGAGAGAGGTCTCTATCCCTGGCAGGTATCTCTGAACAAAGCGCAGACATATT 404
Qy 241 TGTGAGGAGAGCATGCTCTACCCAGTGGGTGATCAGCGGGCTCACTGCTATTCGAAC 300
Db 405 TGTGAGGAGAGCATGCTCTACCCAGTGGGTGATCAGCGGGCTCACTGCTATTCGAAC 464
Qy 301 AGAACATTTGCTACTTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360
Db 465 AGAACATTTGCTACTTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 524
Qy 361 CCAGGAGAGCAACTCTCATTGAACATGTCATCATACATCCCATTTCTCCACCAG 420
Db 525 CCAGGAGAGCAACTCTCATTGAACATGTCATCATACATCCCATTTCTCCACCAG 584
Qy 421 AAACCAATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 480
Db 585 AAACCAATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 644
Qy 481 TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCGTGGTTTTATTTGT 540
Db 645 TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCGTGGTTTTATTTGT 704
Qy 541 ACAACTGCAGGCTGGGGCCGCTTAACCTAAGTGGCGTCTCTACACATCTTGAGGAA 600
Db 705 ACAACTGCAGGCTGGGGCCGCTTAACCTAAGTGGCGTCTCTACACATCTTGAGGAA 764

Qy 601 GTGAATCTGCCTATTTTACCTGGGAGAGTGTGTGGCAGCTCTCTTAACACTAAAGAGG 660
Db 765 GTGAATCTGCCTATTTTACCTGGGAGAGTGTGTGGCAGCTCTCTTAACACTAAAGAGG 824
Qy 661 CCCATCACTGGGAGACCTTTCTTTTGCACAGGTTTCTCTGATGGAGGAGACGCATGT 720
Db 825 CCCATCACTGGGAGACCTTTCTTTTGCACAGGTTTCTCTGATGGAGGAGACGCATGT 884
Qy 721 CAGGAGATTCAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCT 780
Db 885 CAGGAGATTCAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCT 944
Qy 781 GGTGTGACTCTCTGGGTTTGGGCTGTGTGAGCTCGAGAAACAATGTGAGGAAAAGT 840
Db 945 GGTGTGACTCTCTGGGTTTGGGCTGTGTGAGCTCGAGAAACAATGTGAGGAAAAGT 1004
Qy 841 GATCAAGGATCCCTGGGATCTTTCACAGACATTTAGTAAGTCTTCTTGGATCCACGAA 900
Db 1005 GATCAAGGATCCCTGGGATCTTTCACAGACATTTAGTAAGTCTTCTTGGATCCACGAA 1064
Qy 901 CACATCCAACTGGTAACTAA 921
Db 1065 CACATCCAACTGGTAACTAA 1085
XX
XX RESULT 3
XX AAD09331
XX ID AAD09331 standard; cDNA; 1568 BP.
XX AC AAD09331;
XX XX
XX DT 10-SEP-2001 (first entry)
XX XX
XX DE Human protease homologue cDNA #4.
XX KW Human; protease homologue; novel human protein; NHP; therapy;
XX OS pharmacogenomic application; physiological disorder; ss.
XX OS Homo sapiens.
XX PN WO200146407-A1.
XX PD 28-JUN-2001.
XX PF 12-DEC-2000; 2000WO-US33738.
XX PR 23-DEC-1999; 99US-0171566.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX WPI: 2001-408641/43.
XX
XX Polynucleotide encoding novel human protease homologs, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
XX PS Disclosure; Page 31; 32pp; English.
XX
XX The present sequence is a cDNA encoding novel human protein (NHP),
CC known as human protease homologue. NHP shares structural similarity
CC with animal proteases, particularly trypsin-like protease such
CC as oviductin, plasminogen activator and human plasma kallikrein
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
CC pharmacogenomic applications. NHP sequences are useful for identifying
CC agonists, antagonists and modulators and also for producing antibodies
CC useful in diagnosis, drug screening, clinical trial monitoring and in
CC treatment of physiological disorders.
XX
XX Sequence 1568 BP; 452 A; 338 C; 353 G; 423 T; 2 other;

Query Match		99.9%;	Score 920.2;	DB 22;	Length 1568;
Best Local Similarity		100.0%;	Pred. No. 7.5e-285;		
Matches 921;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT	60		
DB	245	ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT	304		
QY	61	TTTGAACRAGGTAATCTGCACTCTTTTCGCTCCCAAGCTCCCAAGCTGTTGGGACAGAGT	120		
DB	305	TTTGAACRAGGTAATCTGCACTCTTTTCGCTCCCAAGCTCCCAAGCTGTTGGGACAGAGT	364		
QY	121	CTGGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTCACTCGCATTTTGGAGGAAGC	180		
DB	365	CTGGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTCACTCGCATTTTGGAGGAAGC	424		
QY	181	CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGCGAAGCATATT	240		
DB	425	CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGCGAAGCATATT	484		
QY	241	TGTGGAGGAAGCATCTCTCACACAGTGGGTGATCAGCGCGCTCACTGCATTGCAAAAC	300		
DB	485	TGTGGAGGAAGCATCTCTCACACAGTGGGTGATCAGCGCGCTCACTGCATTGCAAAAC	544		
QY	301	AGAAACATTGTCTACTTTGAATGTTACTGCTGGAGATGATGACTTAAGCCACAGACAG	360		
DB	545	AGAAACATTGTCTACTTTGAATGTTACTGCTGGAGATGATGACTTAAGCCACAGACAG	604		
QY	361	CCAGGAGAGCAACTCTCACTATTGAACTGTCATCATACATCCACATTTCTCCACCAG	420		
DB	605	CCAGGAGAGCAACTCTCACTATTGAACTGTCATCATACATCCACATTTCTCCACCAG	664		
QY	421	AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC	480		
DB	665	AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC	724		
QY	481	TTTGTGGGCCCCATATGTCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGT	540		
DB	725	TTTGTGGGCCCCATATGTCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGT	784		
QY	541	ACAACCTGCAGGCTGGGGCGCTTAACTGAAGGTGGCGTCTCTCACAAAGTCTTGCAGGAA	600		
DB	785	ACAACCTGCAGGCTGGGGCGCTTAACTGAAGGTGGCGTCTCTCACAAAGTCTTGCAGGAA	844		
QY	601	GTGAATCTGCCATATTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTTAACTACTAAAGAGG	660		
DB	845	GTGAATCTGCCATATTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTTAACTACTAAAGAGG	904		
QY	661	CCCATCAGTGGGAAGACCTTTCTTTTGACAGGTTTTCCTGATGGAGGGAGAGACGCATGT	720		
DB	905	CCCATCAGTGGGAAGACCTTTCTTTTGACAGGTTTTCCTGATGGAGGGAGAGACGCATGT	964		
QY	721	CAGGGAGATTACAGAGGTTCACTCATGTGCCGAATTAAGAAAGGGCGCTTGGACTCTGGCT	780		
DB	965	CAGGGAGATTACAGAGGTTCACTCATGTGCCGAATTAAGAAAGGGCGCTTGGACTCTGGCT	1024		
QY	781	GGTGTGACTTCTCGGGGTTTGGGCTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAAGT	840		
DB	1025	GGTGTGACTTCTCGGGGTTTGGGCTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAAGT	1084		
QY	841	GATCAAGGATCCCTCGGATCTTTCACAGACATTTAGTAAAGTGTCTTCTCGATCCACGAA	900		
DB	1085	GATCAAGGATCCCTCGGATCTTTCACAGACATTTAGTAAAGTGTCTTCTCGATCCACGAA	1144		
QY	901	CACATCCAACTGGTAACATA 921			
DB	1145	CACATCCAACTGGTAACATA 1165			

RESULT 4
AAD09329
ID AAD09329 standard; cDNA; 909 BP.
XX

AC	AAD09329;				
XX	10-SEP-2001 (first entry)				
XX	Human protease homologue cDNA #2.				
XX	Human; protease homologue; novel human protein; NHP; therapy; pharmacogenomic application; physiological disorder; ss.				
OS	Homo sapiens.				
XX	Key	Location/Qualifiers			
FT	CDS	1..909			
FT		/*tag= a			
FT		/product= "Human protease homologue #2"			
FT	unsure	55..57			
FT		/*tag= b			
FT		/note= "Encodes Arg"			
FT	misc_feature	56			
FT		/*tag= c			
FT		/note= "This degenerate base represents a polymorphic site"			
FT	unsure	70..72			
FT		/*tag= d			
FT		/note= "Encodes Ala"			
FT	misc_feature	70			
FT		/*tag= e			
FT		/note= "This degenerate base represents a polymorphic site"			
XX					
PN	W0200146407-A1.				
XX	28-JUN-2001.				
XX					
PF	12-DEC-2000; 2000WO-US33738.				
XX	23-DEC-1999; 99US-0171566.				
XX	(LEXI-) LEXICON GENETICS INC.				
XX	Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;				
XX	WPI; 2001-408641/43.				
DR	P-PSDB; AAE04734.				
XX	Polynucleotide encoding novel human protease homologs, useful for identifying agonist, antagonist or modifiers or for producing antibodies useful in therapeutic, diagnostic and pharmacogenomic applications -				
XX	Disclosure; Page 29; 32pp; English.				
XX	The present sequence is a cDNA encoding novel human protein (NHP), known as human protease homologue. NHP shares structural similarity with animal proteases, particularly trypsin-like protease such as oviductin, plasminogen activator and human plasma kallikrein precursor. NHP and its cDNA are useful in therapeutic, diagnostic and pharmacogenomic applications. NHP sequences are useful for identifying agonists, antagonists and modulators and also for producing antibodies useful in diagnosis, drug screening, clinical trial monitoring and in treatment of physiological disorders.				
XX	Sequence 909 BP; 237 A; 193 C; 238 G; 239 T; 2 other;				
Query Match		98.6%;	Score 908.2;	DB 22;	Length 909;
Best Local Similarity		100.0%;	Pred. No. 4e-281;		
Matches 909;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	13	ATGCTTATAAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTTTGAACRAGGT	72		
DB	1	ATGCTTATAAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTTTGAACRAGGT	60		
QY	73	AAATCTGCARCTCTTTCGCTCCCAAGCTCCCAAGCTGTCGGCAGAGTCTGGTTAAGGTA	132		


```
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..495
FT /*tag= a
FT /product= "Human protease homologue #3"
XX
XX WO200146407-A1.
XX
XX 28-JUN-2001.
XX
XX 12-DEC-2000; 2000WO-US33738.
XX
XX 23-DEC-1999; 99US-0171566.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX
XX WPI; 2001-408641/43.
XX
XX P-PSDB; AAE04735.
XX
XX Polynucleotide encoding novel human protease homologs, useful for
XX identifying agonist, antagonist or modifiers or for producing
XX antibodies useful in therapeutic, diagnostic and pharmacogenomic
XX applications -
XX
XX Disclosure; Page 30; 32pp; English.
XX
XX The present sequence is a cDNA encoding novel human protein (NHP),
XX known as human protease homologue. NHP shares structural similarity
XX with animal proteases, particularly trypsin-like protease such
XX as oviductin, plasminogen activator and human plasma kallikrein
XX precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
XX pharmacogenomic applications. NHP sequences are useful for identifying
XX agonists, antagonists and modulators and also for producing antibodies
XX useful in diagnosis, drug screening, clinical trial monitoring and in
XX treatment of physiological disorders.
XX
XX Sequence 495 BP; 117 A; 101 C; 147 G; 130 T; 0 other;
XX
XX Query Match 53.7%; Score 495; DB 22; Length 495;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-148;
XX Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 427 ATGGACTATGATATTCGCCCTTTTGAAGATGGCTGGAGCCTTCAATTTGGCCACTTTGTG 486
DB 1 ATGGACTATGATATTCGCCCTTTTGAAGATGGCTGGAGCCTTCAATTTGGCCACTTTGTG 60
QY 487 GGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGCTGCTTTTATTCTACAAC 546
DB 61 GGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGCTGCTTTTATTCTACAAC 120
QY 547 GCAGGCTGGGGCGGCTTAAGTGGCGTCTCTCAAGCTCTTGAGGAGTGAAT 606
DB 121 GCAGGCTGGGGCGGCTTAAGTGGCGTCTCTCAAGCTCTTGAGGAGTGAAT 180
QY 607 CTGCCATTTTACCTGGGAAGAGTGTGGGAGCTCTGTTAACAATAAGAGGCCCATC 666
DB 181 CTGCCATTTTACCTGGGAAGAGTGTGGGAGCTCTGTTAACAATAAGAGGCCCATC 240
QY 667 AGTGGGAAGACCTTTCTTTCACAGCTTTTCTGATGGAGGGAGACCGCATCTCAGGA 726
DB 241 AGTGGGAAGACCTTTCTTTCACAGCTTTTCTGATGGAGGGAGACCGCATCTCAGGA 300
QY 727 GATTACGAGGTTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG 786
DB 301 GATTACGAGGTTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG 360
QY 787 ACTTCTCGGGGTTTGGGCTGTGGTGTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAA 846
DB 361 ACTTCTCGGGGTTTGGGCTGTGGTGTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAA 420
```

```
QY 847 GGATCCCTCGGATCTTCACAGACATTAGTAAAGTCTTCTCCGTGATCCACGAACATC 906
DB 421 GGATCCCTCGGATCTTCACAGACATTAGTAAAGTCTTCTCCGTGATCCACGAACATC 480
QY 907 CAACTGGTAACTAA 921
DB 481 CAACTGGTAACTAA 495
RESULT 8
AAC05816
ID AAC05816 standard; cDNA; 324 BP.
XX
XX AAC05816;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 9891.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 9891; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
XX Sequence 324 BP; 91 A; 61 C; 74 G; 94 T; 4 other;
```

```
Query Match 28.2%; Score 260; DB 21; Length 324;
Best Local Similarity 95.8%; Pred. No. 4.9e-73;
Matches 272; Conservative 5; Mismatches 5; Indels 2; Gaps 1;
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGCAACAGCTGATTTTACTACTAGGAATAGTCTTT 60
DB 41 ATGAGTCTCAAAATGCTTATAAGCAGCAACAGCTGATTTTACTACTAGGAATAGTCTTT 100
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTTTCGCTCCCCAAAGCTCCCAAGTT--GTGGGCGA 118
DB 101 TTTGAACRAGGTAAATCTGCARCTCTTTTCGCTCCCCAAAGCTCCCAAGTTTGTGGCGA 160
```

QY 119 GCTGTTAAGGTACAGCCTTGGAAATATTTAAACATTTTCAGTCGCATTTCTTGGAGAA 178
|||||
Db 161 GCTGTTAAGGTACAGCCTTGGAAATATTTAAACATTTTCAGTCGCATTTCTTGGAGAA 220
|||||
QY 179 GCGAAGTGGAGAGGTTTCCATCCCTGCGAGGTATCTCTGAACAAAGGAGAGCATA 238
|||||
Db 221 GCGAAGTGGAGAGGTTTCCATCCCTGCGAGGTATCTCTGAACAAAGGAGAGCATA 280
|||||
QY 239 TTTGTTGGAGGAGCATCTCTCACCACAGTGGGTGATCACGGCG 282
|||||
Db 281 TTTGTTGGAGGAGCATCTCTCACCACAGTGGGTGATCACGGCG 324
|||||
RESULT 9
AAS88433
ID AAS88433 standard; cDNA; 2742 BP.
XX
AC AAS88433;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #24237.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
DR P-PSDB; ABG24246.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 24237; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 2742 BP; 835 A; 561 C; 664 G; 682 T; 0 other;
Query Match 17.7%; Score 163.2; DB 23; Length 2742;
Best Local Similarity 51.7%; Pred. No. 2e-41;
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;
QY 148 TTTAACTTTTCAGTCGCATTTCTTGGAGGAAGCCAAAGTGGAGAGGGTTCCTATCCCTGG 207
|||||
Db 73 TCTAGATTTCTCTAGAAATAGTAGTTGGAGAAATTCACAGTACTGGACATCCATGG 132
|||||
QY 208 CAGGTATCTCTGAACAAAGGAGCAAGCATATTTTGGAGGAAGCATCTCTCACACAG 267
|||||
Db 133 CAGGTCTCCCTAAAATCAGATGAGCACCACCTCTCTGGAGGAAGCTTGAATCAAGAAGAT 192
|||||
QY 268 TGGGTGATCAGCGGGCTCAGTCTGATTCGCAAAACAGA-----ACATTTGTGCTACTTTG 321
|||||
Db 193 CGGGTTGTTACAGCAGCACACTGGCTGGACAGCCCTCAGTGAGAGCAGCTGAAGAATATA 252
|||||
QY 322 AATGTTACTGCTGGAGAGTATGACTTAAGCCAGCAGACACCCAGGAGAGCAAACTCTCACT 381
|||||
Db 253 ACTGTGACTTCTGGGAGTACAGCCTCTTTCAAGAGGATAAGCAAGACAGAAATATTCCT 312
|||||
QY 382 ATTGAAACTGTCAATCATATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATT 441
|||||
Db 313 GTCTCAAAAATTAATACCCATCTCAATACACAGCCGCTGAATATATGAGTCTGATATT 372
|||||
QY 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTTGGGGGCCCATATGTCTT 501
|||||
Db 373 GCAGTGTGTATCTAAAACACAAAGTCAAGTTTGGAAATGCTGTTACGCCAATCTGTCTT 432
|||||
QY 502 CCAGAGCTGCGGGAGCAATTTGAGCTGTTTTTATTTCTACAACCTGCAGGAGTGGGGCCGC 561
|||||
Db 433 CCTGACAGCGATGATAAAGTTGAACCCAGGAATCTTTCTCTATCCAGTGGATGGGGCAAG 492
|||||
QY 562 TTAAGTGAAGTGGCGTCTCTCACAGTCTTGCAGGAAGTGAATCTGCCCTATTTTGACC 621
|||||
Db 493 ATTTCCAAACATCAGATATATCAAAATGCTCTACAAGAAATGGAATTTCCCATCATGGAT 552
|||||
QY 622 TGGGAAGAGTGTGGGAGCTCTGTACACTAAGAGGCCCATCAGTGGGAAGACCTTT 581
|||||
Db 553 GACAGAGCGTGTAACTGTCTCAAGAGCATGAACCTCCCTCCCTGGGAAGGACCATG 612
|||||
QY 682 CTTTGCACAGGTTTTCTGTAGTGGAGGAGAGACGCATCTCAGGGAGATTCAGGAGGTTCA 741
|||||
Db 613 CTGTGCTGCTGCTTCCCTGATTGGGAATGACGCCTCCAGGGGGACTCTGGAGGACCA 672
|||||
QY 742 CTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGTGTGACTTCTCTGGGGTTTG 801
|||||
Db 673 CTGGTTTGTAGAAGAGGTGGTGAATCTGGATTTCTTGTGGGATAACTTCTCTGGTAGCT 732
|||||
QY 802 GGCTGTGCTCAGGCTGGAGAAACAAATGTGAGGAAAAGTGATCAAGGATCCCTGGGATC 861
|||||
Db 733 GGTGTGCTGGAGGTTCAAGTTCCCGTAAGAAACAAACCATGTGAAGGCATCACTTTGGCAT 792
|||||
QY 862 TTCACAGACATTAAGTAAAGTGTCTTCTCGGATCCACGAACACATCCCAACCTGGT 915
|||||
Db 793 TTCCTCAAGTCTGTAGTTGATGATTTTATCACTCAAAACCTGTTTCACAGGT 846
|||||
RESULT 10
AAS87759/c
ID AAS87759 standard; cDNA; 3309 BP.
XX
AC AAS87759;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #23563.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX

```
OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG23572.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 1; SEQ ID No 23563; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 3309 BP; 789 A; 783 C; 683 G; 1054 T; 0 other;
XX Query Match 17.7%; Score 163.2; DB 23; Length 3309;
XX Best Local Similarity 51.7%; Pred. No. 2.2e-41;
XX Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;
XX 148 TTAAACATTTTCAGTCCGATCTTGGAGGAGCCCAAGTGGAGAGGGTTCCTATCCCTGG 207
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 3237 TCTAGATTCTTCTAGATTAGTAGTTGGAGAAATTCACAGTGAAGTCCCATGG 3178
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 208 CAGGTATCTCTGAACAAAGGCAAGCATATTTGGAGGAGCAAGCTCGTCCACACAG 267
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 3177 CAGGTCTCCCTAAATCAGATGAGCACCACTCTGTGGAGGAAGCTTGATTCAAGAAGAT 3118
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 268 TGGGTGATCAGCGGGCTCAGTTCGATTCGAAACAGA-----AACATTGTGTCTACTTTG 321
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 3117 CGGGTTCTACAGCAGCACACTGCTCGACAGCCTCAGTGAGAGCAGCTGAAGAATATA 3058
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 322 AATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACT 381
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 3057 ACTGTGACTTCTGGGAGTACAGCCTCTTTTCAGAAAGTAAAGCAAGCAAGATATTCCT 2998
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 382 ATTGAACACTGTCATACATCCACATTTCTCCACCAGAAACCAATGGACTATGATATT 441
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 2997 GTCTCAAAAATTAATACCATCTCTGAAATACAAAGCCGCTGAATATATGAGTCTGATATT 2938
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

442 GCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACCTTTGTGGGGCCCATATGTCTT 501
2937 GCACTGCTGATCTATAAACACAAAGTCAAGTTTGGAAATGCTGTTCAGCCAAATCTGCTT 2878
502 CCAGAGCTCGGGAGCAATTTGAGGCTGGTATTATTGTACAACTGCAGGCTGGGGCCGC 561
2877 CCTGACAGCGATGATAAAGTTGAACAGGAATCTTTTCTTATCCAGTGGATGGGCAAG 2818
562 TTAACCTGAAGTGGCGTCTCTCACAAGTCTTGAGGAGGAGACCATGTTCAGGAGATTTCAGG 621
2817 ATTTCCAAAACATCAGAAATATCAATGTCTACAAAGAAATGGAATCTCCATCATGGAT 2758
622 TGGGAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCATGTGGGAAGACCTTT 681
2757 GACAGAGCGTGAATACGTGCTCAAGAGCATGAACCTCCCTCCCTGGGAAGACCATG 2698
682 CTTTGCACAGGTTTCTCTGATGGAGGAGAGACCATGTTCAGGAGAGATTTCAGGAGGTTCA 741
2697 CTGTGCTGCTGCTTCCCTGATTTGGGAATGGACGCTTCCAGAGGGGAGCTCTGGAGGACCA 2638
742 CTATGTGCCGGAATAAGAAAGGCGCTGGACTCTGCTGGTGGTGTGACTTCTCTGGGTTTG 801
2637 CTGGTTGTAGAGAGGTTGGTGAATCTGGATCTTCTCTGGGATAACTTCTCTGGGTAGCT 2578
802 GGCTGTGCTGAGCTGGAGGAGGAGAAACAATGTGAGGAAAGTATCAAGATCCCTTGGGATC 861
2577 GGTGTGCTGGAGGTTTCAGTTCCGTAAGAAACAACCATGTGAAGGCATCACTTGGCATT 2518
862 TTCACACATATTAGTAAAGTGTCTTCTGGATCCAGCAACACATCCCAACTGGT 915
2517 TTCTCCAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2464

RESULT 11
AAS72057
ID AAS72057 standard; cDNA: 4729 BP.
AC AAS72057;
DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #7861.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG07870.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 1; SEQ ID No 7861; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

Qy	862	TTCCACGACATTAGTAAAGTGCTTTCTCGATCCAGACACATCCAACCTGGT	915
Dd	3310	TTCTCCAAGTGCTCGAGTTGATGGATTTTATCACTCAAAACCTCTTCACAGGT	3363
RESULT 12			
AAS74405			
ID	AAS74405 standard; cDNA; 4729 BP.		
XX	XX	XX	XX
AC	AAS74405;		
XX	XX	XX	XX
DT	13-FEB-2002 (first entry)		
XX	XX	XX	XX
DE	DNA encoding novel human diagnostic protein #10209.		
XX	XX	XX	XX
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder;		
XX	XX	XX	XX
OS	Homo sapiens.		

US
Homo sapiens.
XX

PN	WO2001/5067-A2.
XX	
XX	
PD	11-OCT-2001.
XX	
XX	
PF	30-MAR-2001; 2001WO-US08631.

XX
PF 30-MAR-2001; 2001WU-US08631.

PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167

XX

XX

XX
+ +
XXXXX / 77 XXXXXX
XXXXX / 77 XXXXXX

DR WFL, 2001-039302/13.
DR P-PSDB: ABG10218.

Now isolated volumetrically and encoded volumetrically in

PT diagnostics, forensics, gene mapping, identification of mutation

PT biodiversity -

PS Claim 1; SEQ ID No 10209; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (T) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probe

CC and gene mapping, and in recombinant production of (II). The

CC for identifying expressed genes. (I) is useful in gene therapy t

to restore normal activity of (II) or to treat disease states in (II). (II) is useful for generating antibodies against it, detecting (II). (II) is useful for generating antibodies against it, detecting

quantitating a polypeptide in tissue, as molecular weight markers

CC imaging of sites expressing (II). (I) and (II) are useful for tr

CC The polypeptide and polynucleotide sequences have applications in

CC responsible for genetic disorders or other traits to assess blood

CC and to produce other types of data and products dependent on DNA
CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

specification, but was obtained in electronic format directly from the FBI.

XX
XX

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 84

Query Match 17.7%; Score 163.2; DB 23; Length 4729;
Best Local Similarity 51.7%; pred. No. 26e-41;

Matches 400; Conservative 0; Mismatches 368; Indels 6;

QY 148 TTTAACAATTTTCAGTCGCATTCTTGAGGAAGCCAAGTGGAGAAGGGTTCCTATCCCTG

Db 2590 TCTAGATTCTCTCTAGAAATAGTAGTTGGAGAAATTCACAGTGACTGGACATCCATGG 2649
Qy 208 CAGGTATCTCTCAACAAGCAGAGCATATTTGTGGAGAAAGCATGCTCTCACCACAG 267
Db 2650 CAGGTCTCCCTAAATACAGATGAGCACCACCTCTGTGGAGAAAGCTTGATTCAAGAAGT 2709
Qy 268 TGGGTGATCAGCGGGCTCACTGATTCGAAACAGA-----AACATTGTGTCTACTTTG 321
Db 2710 CGGGTTGTACAGCAGCAGCTGCTGGACAGCCTCAGTGAGAGCAGCTGAAGAAATATA 2769
Qy 322 AATGTACTGTGGAGAGATGACTTAAGCCAGACAGACCAGGAGCAAACTCTACT 381
Db 2770 ACTGTGACTTCTGGGAGCTACAGCCTCTTTCAGAGGATTAAGCAAGACAGAAATATCT 2829
Qy 382 ATTGAACATGTCATCATACATCCACATTTCTCCACCAAGAAACCAATGGATATGATTT 441
Db 2830 GTCTCAAAAATTTTACCCTTCCGTAATACACAGCCGTGAATATATAGTCCCTGATATT 2889
Qy 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTT 501
Db 2890 GCAGTCTGTATCTAAACACAAAGTCAAGTTTGGAAATGCTGTTCACCAATCTGTCTT 2949
Qy 502 CCAGAGCTCGCGGAGCAATTTGAGCTGCTTTAATTTGTACAACATGACAGGCTGGGGCCG 561
Db 2950 CCTGACAGCGATGATAAAGTTGAACCCAGGAATCTTCTTCTTATCCAGTGGATGGGCAAG 3009
Qy 562 TTAAGTGAAGTGGGCTCTCTCAAGCTCTGACAGAGTGAATCTGCTCTATTTTGACC 621
Db 3010 ATTTCCAAAACATCAGAAATATTTCAATGTCTTACAGAAATGGAATTTCCCATCATGAT 3069
Qy 622 TGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTT 681
Db 3070 GACAGAGCGTGTAACTGTGCTCAAGAGCATGAACCTCCCTCCCTGGGAAGGACCATG 3129
Qy 682 CTTTCCAGAGCTTTCTCTGATGGAGGAGACGATGTGAGGAGATTCAGGAGTTCA 741
Db 3130 CTGTGTGTGCTTCCCTGTATGGGGAATGACGCTCCAGGGGGACTCTGGAGGACCA 3189
Qy 742 CTCATGTGCGGGAATAAGAAAGGGCCCTGGACTCTGGCTGTGTGCTCTCTCTGGGGTTG 801
Db 3190 CTGGTTGTAGAGAGGTGGTGGATCTGGATCTTGTCTGGGATACCTCTCTGGTAGCT 3249
Qy 802 GCTGTGTGTCAGGCTGGAGAAACAATGTGAGGAAAAGTATCAAGGATCCCTCGGATC 861
Db 3250 GCTTGTGCTGGAGTTGAGTTCCCGTAAGAAACACCACTGTGAAGGCATCACTTGGCAT 3309
Qy 862 TTCACAGACATAGTAAAGTCTTTCTCTGGATCCAGACACATCCAAACTGGT 915
Db 3310 TTCTCCAAAGTGTCTGAGTTGATGATTTATCTACTCAAAACCTGTTCACAGGT 3363

RESULT 13

AA578775
ID AA578775 standard; cDNA; 4729 BP.
XX
AC AA578775;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #14579.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W020175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR P-PSDB; ABG14588.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID No 14579; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Query Match 17.7%; Score 163.2; DB 23; Length 4729;
Best Local Similarity 51.7%; Pred. No. 2.6e-41;
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;
Qy 148 TTAAACATTTTCAGTCGCATCTCTGGAGAGCCCAAGTGGAGAGGTTCTCTATCCCTGG 207
Db 2590 TCTAGATTCTCTCTAGAAATAGTAGTTGGAGAAATTCACAGTCACTGGACATCCATGG 2649
Qy 208 CAGGTATCTCTGAAACAAAGCAGAGCATATTTGTGGAGGAAAGCATGCTCTCACCACAG 267
Db 2650 CAGGTCTCCCTAAATCAGATGAGCACCACCTCTCTGTGGAGGAAGCTTGATTCAAGAAGT 2709
Qy 268 TGGGTGATCAGCGGGCTCACTGCAATTCGAAACAGA-----AACATTGTGTCTACTTTG 321
Db 2710 CGGGTTGTACAGCAGCAGCAGCTGCTGGACAGCCTCAGTGAGAGCAGCTGAAGAAATATA 2769
Qy 322 AATGTACTGTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACT 381
Db 2770 ACTGTGACTTCTGGGGAGTACAGCCTCTTTCAGAGGATTAAGCAAGACAGAAATATCT 2829
Qy 382 ATTGAACATGTCATCATACATCCACATTTCTCCACCAAGAAACCAATGGATATGATTT 441
Db 2830 GTCTCAAAAATTTTACCCTATCCTGAATACACAGCCGTGAATATATAGTCTCTGATATT 2889
Qy 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTT 501
Db 2890 GCAGTCTGTATCTTAAAAACACAAAGTCAAGTTTGGAAATGCTGTTCAGCAGCAATCTGCTT 2949
Qy 502 CCAGAGCTCGCGGAGCAATTTGAGCTGCTGTTTATTTGTACAACTGCAGGCTGGGGCCG 561
Db 2950 CCTGACAGGATGATAAAGTTGAACCCAGGAATCTTCTTATCCAGTGGATGGGCAAG 3009

ID ABQ86158 standard; DNA; 3096 BP.

AC ABQ86158;

DT 10-SEP-2002 (first entry)

DE Novel human gene. SEQ ID 29.

XX Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.

XX Homo sapiens.

OS W0200250105-Al.

PN 27-JUN-2002.

PD 17-DEC-2001; 2001WO-US49232.

PF 19-DEC-2000; 2000US-256710P.

PR 20-DEC-2000; 2000US-257048P.

PR 09-JAN-2001; 2001US-260482P.

PR 30-JAN-2001; 2001US-264922P.

PR 06-FEB-2001; 2001US-266797P.

PR 19-MAR-2001; 2001US-276988P.

PR 04-APR-2001; 2001US-281535P.

PR 08-MAY-2001; 2001US-289622P.

XX (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

PA (GLAX) GLAXO GROUP LTD.

XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;

PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;

XX WPI; 2002-508784/54.

DR P-PSDB; ABP60993.

XX Secreted proteins and polynucleotides useful as vaccines for preventing

PT or treating various diseases e.g. cancer, wounds, atherosclerosis,

PT Parkinson's disease, Alzheimer's disease, infection, autoimmune

PT disorder -

XX Claim 2(a); Page 245-246; 335pp; English.

XX The invention relates to an isolated polypeptide with signal sequences

CC which allow it to be secreted extracellularly or membrane associated.

CC The activity of polypeptides of the invention may be described as,

CC cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, nootropic,

CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid

CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-

CC transduction deficiency, neurological diseases, stroke, angiogenesis,

CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,

CC trachea, thymus, lymph node and muscular system, obesity, anorexia,

CC growth abnormalities, and alleviation of precocious puberty. The

CC sequences given in records ABQ86130-ABQ86184 represent novel human cdna's

CC of the invention.

XX Query Match 17.5%; Score 161.6; DB 24; Length 3096;

XX Best Local Similarity 51.6%; Pred. No. 6.8e-41;

XX Matches 399; Conservative 0; Mismatches 369; Indels 6; Gaps 1;

QY 148 TTAACTTTTTCAGTCGCATCTTCTGGAGAGCCCAAGTGGAGAGGGTTCCTATCCCTGG 207

DB 73 TCTAGATTCTTCTAGAAATTAGTAGTTGGAGAAATTAACAGTGCCTGGAGATCCATGG 132

QY 208 CAGGTATCTCTGAAACAAAGCGAGAGCATATTTTGGAGGAGCATGCTCTCACACAG 267

DB 133 CAGGTCTCCCTAAATCAGATGAGCACCACCTTCTCTGGAGGAGCTTGATTCAAGAAGAT 192

QY 268 TGGGTGATACGGCGGCTCAGTCGATTTGCAACAGA-----ACATTGTGCTACTTTG 321

DB 193 CGGGTTGTTACAGCAGCACACTGCTGGACAGCCTCAGTGAGAGCAGCTGAAGAATATA 252

QY 322 ATGTTACTGCTGGAGAGTATGACTTAAGCCAGAGACAGCCAGGAGAGCAAACTCTCACT 381

DB 253 ACTGTGACTTCTGGGAGTACAGCCTCTTTCAGAGGAGTAAGCAAGACAGAAATATTCCT 312

QY 382 ATTTGAACTGTCTATCATATACATCTTCTCCACCAAGAAACCAATGACATATGATTT 441

DB 313 GTCTCAAAATTTATCCCATCTCTGATACACAGCCGTGAATATATGATGCTGATATT 372

QY 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGCTTT 501

DB 373 GCACGTCTGTATCTATAAACAACAAAGTCAAGTTTGGAAATGCTTTCAGGCAATCTGCTT 432

QY 502 CCAGAGCTGCGGGGCAATTTGAGGCTGTTTATTTTGTACAACTGCAGGCTGGGGCCGC 561

DB 433 CCTGACAGGATGATAAGTGTGAACAGGAATCTTTGCTTATCCAGTGGATGGGCAAG 492

QY 562 TTAAGTGAAGTGGGCTCTCTCACAAAGTCTTTCAGGAGAGTGAATCTGCTATTTTGACC 621

DB 493 ATTTCCAAACATCAGAAATATTTCAATGTCTTACAAGAAATGAACTTCCCATCATGGAT 552

QY 622 TGGGAAGAGTGTGTGGCAGCTCTGTTTACACATAAGAGAGGCCCATCAGTGGGAGACCTTTT 681

DB 553 GACAGAGCGTGTAACTGTCTCAAGAGCATGAACCTCCCTCCCTGGAAGGAGGACCATG 612

QY 682 CTTTCACAGGTTTCTTCATGAGGAGAGAGCATGTGAGGAGATTCAGAGGTTTCA 741

DB 613 CTGTGTGTGGTCTCCCTGATTTGGGAAATGGAGCCTGCGAGGGGAGCTCTGGAGGACCA 672

QY 742 CTCATGTGCGGGAATAAGAAAGGGGCTGAGCTCTGGCTGTGTGCTCTCTCTGGGTTTG 801

DB 673 CTGCTTTGTAGAGAGGTGGTGAATCTGGATTTCTTGGGATTAATCTCTCTGGGTAGCT 732

QY 802 GGCTGTGTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAAGGATTCCTCTGGGATTC 861

DB 733 GGTGTGTGGAGTTTCAGTTCCCTTAAGAAACAACCATGTGAAGGCATCACTTGGCAATT 792

QY 862 TTCACAGACATTAGTAAAGTGTCTTCCGGATCCAGGAACACATCAAACTGGT 915

DB 793 TTCTCCAAAGTGTCTGAGTTGATGATTTTATCTACTCAAAACCTGTTTCACAGAT 846

Search completed: March 25, 2003, 05:11:15
Job time : 285 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 05:11:20 ; Search time 247 Seconds
(without alignments)
2891.552 Million cell updates/sec

Title: US-09-735-713A-1
Perfect score: 921
Sequence: 1 atgagctctcaaaatgcttat.....acatccaaactggtactaa 921

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	920.2	99.9	921	9	US-09-735-713A-1
2	920.2	99.9	1568	9	US-09-735-713A-7
3	908.2	98.6	909	9	US-09-735-713A-3
4	771.2	83.7	1671	10	US-09-888-615-47
5	495	53.7	495	9	US-09-735-713A-5
6	114.6	12.4	2913	10	US-09-888-615-42
7	76.2	8.3	1327	9	US-09-978-295A-170
8	76.2	8.3	1327	9	US-09-978-697-170
9	76.2	8.3	1327	9	US-09-978-192A-170
10	76.2	8.3	1327	9	US-09-999-832A-170
11	76.2	8.3	1327	9	US-09-978-189-170
12	76.2	8.3	1327	9	US-09-978-608A-170
13	76.2	8.3	1327	9	US-09-978-191A-170
14	76.2	8.3	1327	9	US-09-978-403A-170
15	76.2	8.3	1327	9	US-09-978-564A-170
16	76.2	8.3	1327	9	US-09-978-585A-170
17	76.2	8.3	1327	9	US-10-017-081A-170
18	76.2	8.3	2409	10	US-09-888-615-54
19	76.2	8.3	3143	9	US-09-978-295A-168

ALIGNMENTS

RESULT 1
US-09-735-713A-1
; Sequence 1, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-1
Query Match 99.9% Score 920.2; DB 9; Length 921;
Best Local Similarity 100.0%; Pred. No. 2.4e-290; Indels 0; Gaps 0;
Matches 921; Conservative 0; Mismatches 0;
QY 1 ATGAGCTCTCAAAATGCTTATAGCAGGACAAAGCTGATTTTACTAGGAATAGTCTTTT 60
|||||
Db 1 ATGAGCTCTCAAAATGCTTATAGCAGGACAAAGCTGATTTTACTAGGAATAGTCTTTT 60
QY 61 TTGTAACRAGGTAATGTCARCTCTTTGCTCCCAAGCTCCAGTTGTGGCAGAGT 120
|||||
Db 61 TTGTAACRAGGTAATGTCARCTCTTTGCTCCCAAGCTCCAGTTGTGGCAGAGT 120
QY 121 CTGGTTAAGGTACAGCTTGGAAATTTATTTACATTTTTCAGTCGCAATCTTGGAGAAGC 180
|||||
Db 121 CTGGTTAAGGTACAGCTTGGAAATTTATTTACATTTTTCAGTCGCAATCTTGGAGAAGC 180

```
QY 181 CAAGTCGAGAAAGGTTCCCTATCCCTGGCAGGATATCTCTGAAACAAGGCGAAGACATATT 240
Db 181 CAAGTCGAGAAAGGTTCCCTATCCCTGGCAGGATATCTCTGAAACAAGGCGAAGACATATT 240
QY 241 TGTGGAGAAAGCATCGTCTCACCACAGCTGGGTGATCAGCGCGGTCACTGCATTGCAAAAC 300
Db 241 TGTGGAGAAAGCATCGTCTCACCACAGCTGGGTGATCAGCGCGGTCACTGCATTGCAAAAC 300
QY 301 AGAAACATTGTGTCTACTTTTGAATGTGTACTGCTGGAGAGTATGACTTTAAAGCCACAGAC 360
Db 301 AGAAACATTGTGTCTACTTTTGAATGTGTACTGCTGGAGAGTATGACTTTAAAGCCACAGAC 360
QY 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTGCATCATACATCCACATTTCCTCCACCAAG 420
Db 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTGCATCATACATCCACATTTCCTCCACCAAG 420
QY 421 AAACCAATGACATGATATATGCCCCTTTTCAAGATGCTGGAGCCTTCCCAATTGGCCAC 480
Db 421 AAACCAATGACATGATATATGCCCCTTTTGAAGATGGCTGGAGCCTTCCCAATTGGCCAC 480
QY 481 TTTGTGGGGCCCATATGCTTCTCCAGAGCTCGGGGAGCAATTTGAGGCTGGTTTATTGT 540
Db 481 TTTGTGGGGCCCATATGCTTCTCCAGAGCTCGGGGAGCAATTTGAGGCTGGTTTATTGT 540
QY 541 ACAACTCGAGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAGTCTTGCAGGAA 600
Db 541 ACAACTCGAGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAGTCTTGCAGGAA 600
QY 601 GTGAATCTGCCTATTTTGACCTGGGAGAGTGTGTGCAGCTCTGTAAACATAAGAGG 660
Db 601 GTGAATCTGCCTATTTTGACCTGGGAGAGTGTGTGCAGCTCTGTAAACATAAGAGG 660
QY 661 CCCATCAGTGGGAGACCTTTCTTGACAGAGTCTTCTCATGAGGAGGAGACGCATGT 720
Db 661 CCCATCAGTGGGAGACCTTTCTTGACAGAGTCTTCTCATGAGGAGGAGACGCATGT 720
QY 721 CAGGAGATTACAGAGGTTCACTCATGTGCGGAAATAGTAAAGTGGCTTCCCTGGATCTGGCT 780
Db 721 CAGGAGATTACAGAGGTTCACTCATGTGCGGAAATAGTAAAGTGGCTTCCCTGGATCTGGCT 780
QY 781 GGTGTGACTTCCCTGGGGTTTGGGCTGTGTCGAGGCTGGAGAAACATGTGAGGAAAGT 840
Db 781 GGTGTGACTTCCCTGGGGTTTGGGCTGTGTCGAGGCTGGAGAAACATGTGAGGAAAGT 840
QY 841 GATCAAGATCCCTCGGATCTTCACAGACATTAGTAAAGTGGCTTCCCTGGATCTGGCT 900
Db 841 GATCAAGATCCCTCGGATCTTCACAGACATTAGTAAAGTGGCTTCCCTGGATCTGGCT 900
QY 901 CACATCCAAACTGGTAACTAA 921
Db 901 CACATCCAAACTGGTAACTAA 921
```

RESULT 2

```
US-09-735-713a-7
; Sequence 7, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abulin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 7
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713a-7

Query Match          99.9%; Score 920.2; DB 9; Length 1568;
Best Local Similarity 100.0%; Pred. No. 3.3e-290;
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
Db 245 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 304
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTTGCTCCCCAAAAGCTCCCACTTTGTGGGCAGGT 120
Db 305 TTTGAACRAGGTAAATCTGCARCTCTTTGCTCCCCAAAAGCTCCCACTTTGTGGGCAGGT 364
QY 121 CTGTTTAAGGTACAGCCTTGGNAATTAATTTTAACATTTTTCAGTCGCAATTTTGGAGGAAGC 180
Db 365 CTGTTTAAGGTACAGCCTTGGNAATTAATTTTAACATTTTTCAGTCGCAATTTTGGAGGAAGC 424
QY 181 CAAGTCGAGAAAGGTTCCCTATCCCTGGCAGGTATCTCTGAAACAAGGCGAAGACATATT 240
Db 425 CAAGTCGAGAAAGGTTCCCTATCCCTGGCAGGTATCTCTGAAACAAGGCGAAGACATATT 484
QY 241 TGTGGAGAAAGCATCGTCTCACCACAGTGGGTGATCAGCGCGGTCACTGCATTGCAAAAC 300
Db 485 TGTGGAGAAAGCATCGTCTCACCACAGTGGGTGATCAGCGCGGTCACTGCATTGCAAAAC 544
QY 301 AGAAACATTGTGTCTACTTTTGAATGTGTACTGCTGGAGAGTATGACTTTAAAGCCACAGAC 360
Db 545 AGAAACATTGTGTCTACTTTTGAATGTGTACTGCTGGAGAGTATGACTTTAAAGCCACAGAC 604
QY 361 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCTATCATATCATCATATTTCTCCACCAG 420
Db 605 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCTATCATATCATCATATTTCTCCACCAG 664
QY 421 AAACCAATGACATATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 480
Db 665 AAACCAATGACATATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 724
QY 481 TTTGTGGGGCCCATATGTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 540
Db 725 TTTGTGGGGCCCATATGTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 784
QY 541 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGCGTCTCTCACAAGTCTTTCACAGAA 600
Db 785 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGCGTCTCTCACAAGTCTTTCACAGAA 844
QY 601 GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 660
Db 845 GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 904
QY 661 CCCATCAGTGGGAGACCTTTCTTTGACAGAGTCTTCCCTGATGGAGGAGAGACGCATGT 720
Db 905 CCCATCAGTGGGAGACCTTTCTTTGACAGAGTCTTCCCTGATGGAGGAGAGACGCATGT 964
QY 721 CAGGAGATTACAGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCCTGGACTCTGGCT 780
Db 965 CAGGAGATTACAGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCCTGGACTCTGGCT 1024
QY 781 GGTGTGACTTCCCTGGGGTTTGGGCTGTGTCGAGGCTGGAGAAACATGTGAGGAAAGT 840
Db 1025 GGTGTGACTTCCCTGGGGTTTGGGCTGTGTCGAGGCTGGAGAAACATGTGAGGAAAGT 1084
QY 841 GATCAAGATCCCTCGGATCTTCACAGACATTAGTAAAGTGGCTTCCCTGGATCCACGAA 900
Db 1085 GATCAAGATCCCTCGGATCTTCACAGACATTAGTAAAGTGGCTTCCCTGGATCCACGAA 1144
QY 901 CACATCCAAACTGGTAACTAA 921
Db 1145 CACATCCAAACTGGTAACTAA 1165
```

```
RESULT 3
US-09-735-713A-3
; Sequence 3, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedlich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-3

Query Match      98.6%; Score 908.2; DB 9; Length 909;
Best Local Similarity 100.0%; Pred. No. 2e-286;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ATGCTTAAACGACGACAAAGCTGATTTACTACTAGGAATAGTCTTTTGAACRAGGT 72
DB 1 ATGCTTAAACGACGACAAAGCTGATTTACTACTAGGAATAGTCTTTTGAACRAGGT 60
QY 73 AAATCTGCARCTCTTTTCGCTCCCCAAAGCTCCCAAGTGTGTGGCGAGAGTCTGTTAAAGTA 132
DB 61 AAATCTGCARCTCTTTTCGCTCCCCAAAGCTCCCAAGTGTGTGGCGAGAGTCTGTTAAAGTA 120
QY 133 CAGCCTTGGAAATTTTAAACATTTTTCAGTCGCATTTCTTGAGGAGCAAGTGGAGAG 192
DB 121 CAGCCTTGGAAATTTTAAACATTTTTCAGTCGCATTTCTTGAGGAGCAAGTGGAGAG 180
QY 193 GTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGGACGAGCATATTTGTGGAGGAGC 252
DB 181 GTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGGACGAGCATATTTGTGGAGGAGC 240
QY 253 ATGCTCTACACAGTGGGTGATCAGCGCGCTCACTGCAATTCGCAACAGAAACATTTGTG 312
DB 241 ATGCTCTACACAGTGGGTGATCAGCGCGCTCACTGCAATTCGCAACAGAAACATTTGTG 300
QY 313 TCTACTTTGAATGTTACTGCTGGAGATGATGACTTAAGCCAGACAGACCCAGGAGACAA 372
DB 301 TCTACTTTGAATGTTACTGCTGGAGATGATGACTTAAGCCAGACAGACCCAGGAGACAA 360
QY 373 ACTCTCACTATTGAACCTGTCTATCATATCATATTCATCTCCACCAAGAACCAATGAC 432
DB 361 ACTCTCACTATTGAACCTGTCTATCATATCATATTCATCTCCACCAAGAACCAATGAC 420
QY 433 TATGATATGCCCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCC 492
DB 421 TATGATATGCCCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCC 480
QY 493 ATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTGTATTTTGTACAACTGACAGGC 552
DB 481 ATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTGTATTTTGTACAACTGACAGGC 540
QY 553 TGGGGCCGCTTAACTGAAGTGGCGTCTCTCAACAGTCTTGCAGGAGTGAATCTGCCT 612
DB 541 TGGGGCCGCTTAACTGAAGTGGCGTCTCTCAACAGTCTTGCAGGAGTGAATCTGCCT 600
QY 613 ATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTTAAAGAGGCCCATCATGTTGG 672
```

```
Db 601 ATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTTAAAGAGGCCCATCATGTTGG 660
QY 673 AAGACCTTTTCTTTCACACAGGTTTTCCTGATGAGGGAGAGACGATGTCAAGGAGATTCA 732
Db 661 AAGACCTTTTCTTTCACACAGGTTTTCCTGATGAGGGAGAGACGATGTCAAGGAGATTCA 720
QY 733 GGAGGTTTCACTCATGTCCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCC 792
Db 721 GGAGGTTTCACTCATGTCCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCC 780
QY 793 TGGGTTTGGGCTGTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTGTCAAGGATCC 852
Db 781 TGGGTTTGGGCTGTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTGTCAAGGATCC 840
QY 853 CCTGGATCTTTCACAGACATTTAGTAAGTGTCTTCTCGGATCCAGGAACATCCAAACT 912
Db 841 CCTGGATCTTTCACAGACATTTAGTAAGTGTCTTCTCGGATCCAGGAACATCCAAACT 900
QY 913 GGTAACATAA 921
Db 901 GGTAACATAA 909

RESULT 4
US-09-888-615-47
; Sequence 47, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERRARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-47

Query Match      83.7%; Score 771.2; DB 10; Length 1671;
Best Local Similarity 99.6%; Pred. No. 1.8e-241;
Matches 781; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGAGTCTCAAAATGCTTTAAGCAGGAAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
Db 1 ATGAGTCTCAAAATGCTTTAAGCAGGAAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
QY 61 TTTGAACRAGGTAATCTGCARCTCTTTCGCTCCCCAAAGCTCCAGTGTGGGCGAGAT 120
Db 61 TTTGAACRAGGTAATCTGCARCTCTTTCGCTCCCCAAAGCTCCAGTGTGGGCGAGAT 120
QY 121 CTGTTTAAAGTACAGCCCTTGAATATTTTAAACATTTTTCAGTCGATCTTTGGAGGAAGC 180
Db 121 CTGTTTAAAGTACAGCCCTTGAATATTTTAAACATTTTTCAGTCGATCTTTGGAGGAAGC 180
QY 181 CAAGTGGAGAAAGGTTTCTTATCCCTGGCAGGTATCTCTGAACAAAGGAGCAAGCATATT 240
Db 181 CAAGTGGAGAAAGGTTTCTTATCCCTGGCAGGTATCTCTGAACAAAGGAGCAAGCATATT 240
QY 241 TGTGGAGAAACATCGTCTCACACAGTGGGTGATCAGCGGGCTCAGTGCATTTGCAAAAC 300
Db 241 TGTGGAGAAACATCGTCTCACACAGTGGGTGATCAGCGGGCTCAGTGCATTTGCAAAAC 300
```

```
QY 301 AGAACAATTGTGTCTACTTTTGAATGTTACTGCTGGAGAGTATGACTTTAAGCCAGACAGAC 360
Db 301 AGAACAATTGTGTCTACTTTTGAATGTTACTGCTGGAGAGTATGACTTTAAGCCAGACAGAC 360
QY 361 CCAGGAGACAAACTCTCACTATGAAACTGTCATCATACATCCACATTTCTCCACCAC 420
Db 361 CCAGGAGACAAACTCTCACTATGAAACTGTCATCATACATCCACATTTCTCCACCAC 420
QY 421 AAACCAATGGACTATGATATGCCCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 480
Db 421 AAACCAATGGACTATGATATGCCCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 480
QY 481 TTTGTGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCGCTGGTTTTATTGT 540
Db 481 TTTGTGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCGCTGGTTTTATTGT 540
QY 541 ACAACTGCAGGCTGGGGCCGCTTAAGTGAAGTGGCGTCTCTCAAGTCTTGAGGAA 600
Db 541 ACAACTGCAGGCTGGGGCCGCTTAAGTGAAGTGGCGTCTCTCAAGTCTTGAGGAA 600
QY 601 GTGAATCTGCCTATTATTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
Db 601 GTGAATCTGCCTATTATTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
QY 661 CCCATCAGTGGGAAGACCTTTCTTGCACAGCTTTTCCGTGATGGAGGGAGAGACGCATGT 720
Db 661 CCCATCAGTGGGAAGACCTTTCTTGCACAGCTTTTCCGTGATGGAGGGAGAGACGCATGT 720
QY 721 CAGGAGATTCAGGAGGTTCACTCATGTGCGGGAATAAGAAAGGGCCCTTGGACTCTGGC 779
Db 721 CAGGAGATTCAGGAGGTTCACTCATGTGCGGGAATAAGAAAGGGCCCTTGGACTCTGGC 780
QY 780 TGGT 783
Db 781 TGGT 784

RESULT 5
US-09-735-713A-5
; Sequence 5, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT FILING DATE: 2000-12-12
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-5

Query Match 53.7%; Score 495; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.7e-151;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 427 ATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTG 486
Db 1 ATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTG 60
QY 487 GGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTGTACAAC 546
```

```
Db 61 GGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTGTACAAC 120
QY 547 GCAGGCTGGGGCCGCTTAAGTGAAGTGGCGTCTCTCAACAAGTCTTGCGAGGAAGTGAAT 606
Db 121 GCAGGCTGGGGCCGCTTAAGTGAAGTGGCGTCTCTCAACAAGTCTTGCGAGGAAGTGAAT 180
QY 607 CTGCTTATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTCTTAACACTAAAGAGGGCCCATC 666
Db 181 CTGCTTATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTCTTAACACTAAAGAGGGCCCATC 240
QY 667 AGTGGGAAGACCTTTCTTTGCAAGGTTTCTGATGGAGGAGAGACCATGTCAGGGA 726
Db 241 AGTGGGAAGACCTTTCTTTGCAAGGTTTCTGATGGAGGAGAGACCATGTCAGGGA 300
QY 727 GATTGAGGAGGTTCACTCATGTGCGGGAATAAGAAAGGGCCCTTGGACTCTGGCTGTGTG 786
Db 301 GATTGAGGAGGTTCACTCATGTGCGGGAATAAGAAAGGGCCCTTGGACTCTGGCTGTGTG 360
QY 787 ACTTCTCTGGGTTTGGGCTGTGTGCGAGCTGGAGAAACAATGTGAGGAAAAAGTGTATCAA 846
Db 361 ACTTCTCTGGGTTTGGGCTGTGTGCGAGCTGGAGAAACAATGTGAGGAAAAAGTGTATCAA 420
QY 847 GGATCCCCCTGGGATCTTTCACAGACATTAAGTAAAGTCTTCTTGGATCCACGACACATC 906
Db 421 GGATCCCCCTGGGATCTTTCACAGACATTAAGTAAAGTCTTCTTGGATCCACGACACATC 480
QY 907 CAAACTGGTAACTAA 921
Db 481 CAAACTGGTAACTAA 495

RESULT 6
US-09-888-615-42
; Sequence 42, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 2913
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-42

Query Match 12.4%; Score 114.6; DB 10; Length 2913;
Best Local Similarity 52.4%; Pred. No. 1.3e-26;
Matches 252; Conservative 0; Mismatches 229; Indels 0; Gaps 0;
```

```
QY 435 TGATATTTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGCCCAT 494
Db 9 TGATATTTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTTCAGCCAA 68
QY 495 ATGTCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTGTACAACTGTCAGGCTG 554
Db 69 CTGTCTTCTGACAGGATGATAAAGTTGAACCAAGAAATCTTGTCTTATCAGTGGATG 128
QY 555 GGGCCGCTTAACTGAAGTGGCGTCTCTCAACAAGTCTTTCAGGAAGTGAATCTGCCTAT 614
Db 129 GGGCAAGATTTCCAAAACATCAGATATTAATTAATGTCTTACAAAGAAATGGAATCTCCCAT 188
QY 615 TTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAA 674
```

Db 189 CATGGATGACAGCGCTGTAATACCTGTGCTCAAGACCATGAACCTCCCTCCCTGGGAAG 248
QY 675 GACCTTTCTTTCACAGGTTTTCCTGATGAGGAGGAGAGCATGTGACGGGAGATTTCAGG 734
Db 249 GACCATGCTGTGTGCTGGCTTCCCTGATTTGGGAAATGGAGCGCTGCCAGGGGACTCTGG 308
QY 735 AGTTTCACATCATGTGCGCGGAATAAGAAAGGCGCTGGACTCTGGCTGGTGTGACTTCCCTG 794
Db 309 AGGACCATGCTGTTTGTAGAAAGGTGGTGGATCTGGATCTTGTGGGATAACTTCCCTG 368
QY 795 GGGTTTGGCGCTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGATGCAAGGATCCCC 854
Db 369 GGTAGCTGGTGTGCTGGAGTTCACTCCCGTAAGAAACAACCATGTGGAAGGCATCACT 428
QY 855 TGGGATCTTCACAGACATTAAGTAAAGTCTTCCCTGGATCCAGGACACATCCAAACTGG 914
Db 429 TGGCATTTTCTCCAAAGTCTGTAGTTGATGATTTTATCACTCAAAACCTGTTTCACAGG 488
QY 915 T 915
Db 489 T 489
RESULT 7
US-09-978-295A-170
; Sequence 170, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366

[illegible]

[illegible]

RESULT 10

[illegible]

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C22
CURRENT APPLICATION NUMBER: US/09/978, 608A
CURRENT FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 624
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 170
LENGTH: 1327
TYPE: DNA
ORGANISM: Homo sapiens
US-09-978-608A-170

Query Match 8.3%; Score 76.2; DB 9; Length 1327;
Best Local Similarity 47.2%; Pred. No. 2.8e-14;
Matches 311; Conservative 0; Mismatches 333; Indels 15; Gaps 2;

Qy 157 TTCAGTCGCATCTTGTGAGGAAGCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCT 216
Db 531 TCCAGCCGATGTGTGGAGCTGTCTCCCGAGGGTGAGTGGCCATGGCAGGCCAGC 590
Qy 217 CTGAACAAGGAGAGCATATTTGTGGAGGAAGCATCTCTACACACAGTGGGTGATC 276
Db 591 CTCAGGTTTCGGGTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATA 650
Qy 277 ACGCGCTCCTCACTGCAATGCAACAGAACATCTGTCTACTTTGATGTACTGCTGGA 336
Db 651 ACAGTCGCCACTCTTCCAGGAGGAGCATGCGCTCCACGGTGTGTGGACCGTGTTC 710
Qy 337 GAGTATGACTTAAAGCCAGACAGACCCAGGA---GAGCAAACTCTCACTATTGAAACTGTC 393
Db 711 CTGGCAAGGTGTGGCAGAACTCGCGCTGGCTGGAGAGGTGCTTCAAGGTGACGCCG 770
Qy 394 ATCATATCATCATTTCTCCACCAAGAACCAATGAGTACATGATATGCGCTTTTGAAG 453
Db 771 CTGCTCTGCAACCGGTACCACAGAGAGAGACAGCCATGACTACGACGTGGCGCTGCTGAG 830
Qy 454 ATGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGGCCATATGCTTTCAGAGAGCTCGG 513
Db 831 CTCGACCAACCGGTGGTGGCTCGGCCCGCGCTGGCCCGCTGCTGCGCGCGCTCC 890
Qy 514 GAGCAATTTAGGCTGTTTTTATTTGACAACTGACAGCTGGGGCCGCTTAACTGAAGGT 573
Db 891 CACTTCTCGAGCCGCGCTGCACTGCTGATACGGCTGGGGCGCTTTCGCGAGGGC 950
Qy 574 GGCCTCTCTCACAAGTCTTGCAGGAAGTGAATCTGCTTATTTGACCTGGGAGAGTGT 633
Db 951 GGCCCCATCAGCAACGCTCTGCAAAAGTGGATGTGCGAGTTGATCCACAGGACCTGTGC 1010
Qy 634 GTGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTTGCACAGGT 693
Db 1011 AGCAGGCTTATCGCTACAGGTGACCCACGCA-----TGCTGTGTGCGCGC 1058
Qy 694 TTTCCCTGATGAGGAGAGACGATGTCAGGAGATTCAGAGGTTCACTCATGTGCGCG 753
Db 1059 TACGCAAGGCGAAGAGATGCTGTGACGGTGACTCAGGTGCTGCGCTGGTGTGCAAG 1118
Qy 754 AATAAGAAAGGCGCTGACTCTGGCTGGTGTGACTTCCTCTGGGTTTGGGCTGTGGTCG 812
Db 1119 GCATCAGTGGCGCTGCTGCTGGCGGGCTGTGTCAGCTGGGCGCTGTGGCGC 1177

RESULT 13
US-09-978-191A-170
; Sequence 170, Application US/09978191A
; Publication No. US20030050239A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C4
CURRENT APPLICATION NUMBER: US/09/978, 191A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC25
CURRENT APPLICATION NUMBER: US/09/978,564A
CURRENT FILING DATE: 2001-10-16
PRIOR FILING DATE: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366

[illegible]

Search completed: March 25, 2003, 06:51:50
Job time : 254 secs

THIS PAGE BLANK (USPTO)


```
Db 1751 GACGATCAACAGGCTCGGAGCTTCTCTGCTGAGCAGGGATTTGGCTGGTTCGGCCGCC 1810
Qy 286 CACTGCAATGCAACAGCAACATTTGTCTACTTTTGAATGTTACTGTGGAG-----AG 339
Db 1811 CACTGCTGTACGGGAGAAATATGAGCGCTTAAGTGAAGCAGTCTAGGCTGCAT 1870
Qy 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTGCATATA 399
Db 1871 ATGGCATCAAAATCTGACTTCTCTCAGATAGAAACTAGTTGATTGACCAAAATGTCTATA 1930
Qy 400 CATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTTGCCCTTTTCAAGATGCT 459
Db 1931 AACCCACA---CTACAATAACGAGAAAGAAACAATGACATTTGCCATGATCATCTTGAA 1987
Qy 460 GGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATGCTTCCAGAGCTCGGGGAGCAA 519
Db 1988 ATGAAGTGAACACTACACAGATTATATACAGCTATTGTTTACCAGAAAGAAATCAAGTT 2047
Qy 520 TTTGAGGCTGGTTTATTTGTACAACTGCAGCTGGGCCCTTTAACTGAAGGTGGCGTC 579
Db 2048 TTTCCCCCAGGAAGAAATTTGTTCTATTGCTGGCTGGGGGCGCACTTATATATCAAGGTTCT 2107
Qy 580 CTCCTACAAGCTTTGCAGGAAGTGAATCTGCTATTGTTGACCTGGGAAGAGTGTGSCA 639
Db 2108 ACTGCAGACGTACTGCAAGAAGCTGACCTTCCCTTCTATCAATGAGAATGTCAACAA 2167
Qy 640 GCTCTGTTTAACTAAAGAGCCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTTCCCT 699
Db 2168 CAGATGCCAGAAATAA-----CATTACGGAAATAATGTTGTGTCAGGCTATGAA 2218
Qy 700 GATGGAGGAGAGCGCATGTCAGGAGATTCAGGAGTTTCACATCATGTCGCCGAATAAG 759
Db 2219 GCAGGAGGGTGTAGATCTTGTTCAGGGGGATTTCAGGCGGACCATCATGTGCGCAAGAAAC 2278
Qy 760 AAAGGGGCTGACTCTGGCTGGTGTGACTTCTCTGGGCTTTGGGCTGTG 808
Db 2279 AACAG---ATGGCTCTGGCTGGCTGACGTCAATTTGGATATCAATGTG 2324

RESULT 2
PCT-US94-00616-1
; Sequence 1, Application PC/TUS9400616
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF SEQUENCES: 33
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00616
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US94-00616-1
```

```
Query Match 7.7% Score 71; DB 5; Length 2581;
Best Local Similarity 49.08; Pred. No. 1.9e-13;
Matches 318; Conservative 0; Mismatches 310; Indels 21; Gaps 4;

Qy 166 ATTCTTGGAGGAGCCCAAGTGGAGAGGGTTTCCTATCCCTGGCAGGTATCTCTGAAACAA 225
Db 1691 ATTGTGGAGGAGAGTGACTCCAGAGAGGAGCCTGGGCTTGGTCTGCTCTGATTTTC 1750
Qy 226 AGGCAGAGACATATTTTGGAGGAGACATGCTCTCACCACAGTGGGTGATCAGCGGCGCT 285
```

```
Db 1751 GACGATCAACAGGCTCGGAGCTTCTCTGCTGAGCAGGGATTTGGCTGGTTCGGCCGCC 1810
Qy 286 CACTGCAATGCAACAGCAACATTTGTCTACTTTTGAATGTTACTGTGGAG-----AG 339
Db 1811 CACTGCTGTACGGGAGAAATATGAGCGCTTAAGTGAAGCAGTCTAGGCTGCAT 1870
Qy 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTGCATATA 399
Db 1871 ATGGCATCAAAATCTGACTTCTCTCAGATAGAAACTAGTTGATTGACCAAAATGTCTATA 1930
Qy 400 CATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTTGCCCTTTTGAAGATGCT 459
Db 1931 AACCCACA---CTACAATAACGAGAAAGAAACAATGACATTTGCCATGATCATCTTGAA 1987
Qy 460 GGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATGCTTCCAGAGCTCGGGGAGCAA 519
Db 1988 ATGAAGTGAACACTACACAGATTATATACAGCTATTGTTTACCAGAAAGAAATCAAGTT 2047
Qy 520 TTTGAGGCTGGTTTATTTGTACAACTGCAGCTGGGCCCTTTAACTGAAGGTGGCGTC 579
Db 2048 TTTCCCCCAGGAAGAAATTTGTTCTATTGCTGGCTGGGGGCGCACTTATATATCAAGGTTCT 2107
Qy 580 CTCCTACAAGCTTTGCAGGAAGTGAATCTGCTATTGTTGACCTGGGAAGAGTGTGSCA 639
Db 2108 ACTGCAGACGTACTGCAAGAAGCTGACGTTCCTTCTATCAATGAGAATGTCAACAA 2167
Qy 640 GCTCTGTTTAACTAAAGAGCGCCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTTCCCT 699
Db 2168 CAGATGCCAGAAATAA-----CATTACGGAAATAATGTTGTGTCAGGCTATGAA 2218
Qy 700 GATGGAGGAGAGCGCATGTCAGGAGATTCAGGAGGTTCACATCATGTCGCCGAATAAG 759
Db 2219 GCAGGAGGGTGTAGATCTTGTTCAGGGGGATTTCAGGCGGACCATCATGTGCGCAAGAAAC 2278
Qy 760 AAAGGGGCTGACTCTGGCTGGTGTGACTTCTCTGGGCTTTGGGCTGTG 808
Db 2279 AACAG---ATGGCTCTGGCTGGCTGACGTCAATTTGGATATCAATGTG 2324

RESULT 3
US-09-280-116-107
; Sequence 107, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
; NAME/KEY: misc_feature
; LOCATION: (1)..(796)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-107

Query Match 5.9% Score 54; DB 4; Length 796;
Best Local Similarity 45.98; Pred. No. 4.7e-08;
Matches 220; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

Qy 157 TTCAGTCGCATCTTTGGAGGAGCAAGTGGAGAGGGTTCTCTATCCCTGGCAGGTATCT 216
Db 252 TCCAGCCGCATTTGTTGGTGGAGCTGTGCTCCCGAGGCTGAGTGGCCATGCGAGGCCAGC 311
Qy 217 CTGAACAAGGCAAGACATATTTTGGAGGAAGCATCTCTCACCACAGTGGGTGATC 276
```



```
Db 312 CTCAGGTCGGGGTCGACACATCTGTGGGGGGCCCTCATCGTGCACCGTGGGTGATA 371
Qy 277 ACGGGGCTCACTGATTCGAAACAGAAACATTTGTTGCTACTTTTGAATTTACTCTCTGGA 336
Db 372 ACAGTGGCCCACTGTTCCAGGAGGACAGATGGCTCCACGGTCTGTGGACCGTGTTC 431
Qy 337 GAGTATGACTTAAAGCAGACAGACCCAGGAGAGCAAACTCT--CACTATTGAAACTGTC 393
Db 432 CTGGGCAAGGTGTGGCAGAACTCGCGCTGGCTGGAGAGGTGTCTTCAAGGTGAGCGGC 491
Qy 394 ATCATACATCCACATTTCTCCACCAGAAACCAATGGACTATGATATTTGCCCTTTTGAAG 453
Db 492 CTGCTCTCTGCAACCCGTACCAAGAGAGAGACAGCCATGACTACGAGTGGCGCTGTGCGAG 551
Qy 454 ATGGCTGGAGCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTGGG 513
Db 552 CTCGACCAACCGGTGGTGGCTGGCGCCGCTGGCCCGCTGTGCTGCCCGCGGCTCC 611
Qy 514 GAGCAATTTGAGGCTGTTTATTGTACAACTGCAGGCTGGGCGGCTTAACTGAAGGT 573
Db 612 CACTTCTCGAGCGCGCTGCACCTGCTGGATTAGGGCTGGGGCGCTTGGCGAGGCG 671
Qy 574 GCGTCTCTCACAGTCTTCAGAGAGTGAATCTGCTATTTTGAACCTGGGAGAGTG 632
Db 672 GCGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCANTTTGATCCACAGACCTGTG 730

RESULT 4
US-09-280-116-57
; Sequence 57, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/178965
; CURRENT APPLICATION NUMBER: US/09/280.116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-57

Query Match 5.7%; Score 52.6; DB 4; Length 654;
Best Local Similarity 46.7%; Pred. No. 1.2e-07;
Matches 203; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Qy 154 ATTTTCAGTCGCATCTTTGGAGGAGCCAAAGTGGAGAGGGTTCTATCCCTGGCAGGTA 213
Db 118 ATGCTGAACCAATTTGGTGGGGGGGAGGACACGAGGAGGGCGAGTGGCCCTGGCAAGTC 177
Qy 214 TCTCTGAAACAAAGCAGCAAGCATATTTGTGAGGAACATCGTCTACCACTGGGTG 273
Db 178 AGCATCCAGCGCAACGGAAGCCACTTCTCGGGGGGACGCTCATCGCGGACGAGTGGTGC 237
Qy 274 ATCAGGGGGCTCACTGATGCAACAGAAACATTTGTCTACTTTGAATTTACTGCT 333
Db 238 CTGAGGGTGGCGACTGTCTCGCAACACCTCTGAGAGCTGCCCTGTACAGGTCCTGTGTG 297
Qy 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAACTCTCACTATTGAACACTGTC 393
Db 298 GGGCAAGGCACTAGTGCACCGGGGACCAACGCTATGATGCCGGGTGAGGCAGGTG 357
Qy 394 ATCATACATCCACATTTCTCCCAAGAAACCAATGGACTATGATATTTGCCCTTTTGAAG 453
Db 358 GAGAGCAACCCCTGTACCAAGGACAGCGCTCCA---CGCTGAGGTGGCCCTGGTGGAG 414
Qy 454 ATGGCTGAGCGCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTGGG 513
```

```
Db 415 CTGGAGGCACCAAGTGCCTTCCACCAATTACATCTCCCGTGTGCTGCTGACCCCTCG 474
Qy 514 GAGCAATTTGAGGCTGGTATTATTTGTACAACCTGCAGGCTGGGGCGCTTACTCAAGGT 573
Db 475 GTGATCTTTGAGACGGGCATGAACCTGCTGGGTCACTGGGTGGGGCAGCCCCAGTGAGGAA 534
Qy 574 GCGCTCCTCTCTCACAA 588
Db 535 GACCTCCTGCCCCGAA 549

RESULT 5
US-09-386-653A-1
; Sequence 1, Application US/09386653A
; Patent No. 6458564
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/09/386.653A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-386-653A-1

Query Match 5.7%; Score 52.6; DB 4; Length 1110;
Best Local Similarity 46.7%; Pred. No. 1.7e-07;
Matches 203; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Qy 154 ATTTTCAGTCGCATCTTTGGAGGAGCCAAAGTGGAGAGGGTTCTATCCCTGGCAGGTA 213
Db 126 ATGCTGAACCAATTTGGTGGGGGGCAGGACACGAGGAGGGCGAGTGGCCCTGGCAAGTC 185
Qy 214 TCTCTGAAACAAAGCAGCAAGCATATTTGTGAGGAACATCGTCTACCACTGGGTG 273
Db 186 AGCATCCAGCGCAACGGAAGCCACTTCTCGGGGGGACGCTCATCGCGGAGCAGTGGGTC 245
Qy 274 ATCAGGGGGCTCACTGATGCAACAGAAACATTTGTCTACTTTGAATTTACTGCT 333
Db 246 CTGAGGGTGGCGACTGTCTCCGCAACACCTCTGAGAGCTCCCTGTACCAAGTCTCTGTG 305
Qy 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAACTCTCACTATTGAAACTGTC 393
Db 306 GGGCAAGGCACTAGTGCACCGGGGACCAACGCTATGATGCCGGGTGAGGCAGGTG 365
Qy 394 ATCATACATCCACATTTCTCCCAAGAAACCAATGGACTATGATATTTGCCCTTTTGAAG 453
Db 366 GAGAGCAACCCCTGTACCAAGGACAGCGCTCCA---CGCTGAGGTGGCCCTGGTGGAG 422
Qy 454 ATGGCTGAGCGCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTGGG 513
Db 423 CTGAGGACCACTAGTGCACCGGGGACCAACGCTATGATCTCCCGGTGTGCTGCTGCTGCTG 482
Qy 514 GAGCAATTTGAGGCTGGTATTATTTGTACAACCTGCAGGCTGGGGCGCTTAACTCAAGGT 573
Db 483 GTGATCTTTGAGAGCGGCATGAACCTGCTGGGTCACTGGGTGGGGCAGCCCCAGTGAGGAA 542
Qy 574 GCGCTCCTCTCACAA 588
Db 543 GACCTCCTGCCCCGAA 557

RESULT 6
US-09-386-653A-8
; Sequence 8, Application US/09386653A
```

```
; Patent No. 6458564
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE OF INVENTION: protease T
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/09/386,653A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of
; OTHER INFORMATION: Protease T in a zymogen activation vector
US-09-386-653A-8

Query Match          5.6%; Score 51.8; DB 4; Length 1130;
Best Local Similarity 46.8%; Pred. No. 3.2e-07;
Matches 198; Conservative 0; Mismatches 222; Indels 3; Gaps 1;

QY 166 ATCTTTGGAGAACCAAGTGGAGAGGGTTCCCTATCCCTGGCAGGTATCTCTGAACAA 225
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 ATCGTTGGGGGCTATGCTCTAGAGAGGGGAGTGGCCCTGGCAAGTCAGCATCCAGCGC 225
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 226 AGCAGAAGCATATTTGTGGAGGAAGCATCGTCTACACAGATGGGTGATCAGCGGGGCT 285
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 AACGGAAGCCACATTTCTCGGGGGCAGCCTCATCGCGGAGCAGTGGGTCTCTGACGGCTGC 285
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 286 CACTGCATTGCAACAGAACATTTGCTCTACTTTGAATGTTACTCTGCTGAGAGTATGAC 345
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 CACTGCTTCGCAACACCTCTGAGACGTCCTGTACACAGTCTCTGCTGGGGGCAAGCCAG 345
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 346 TTAAGCAGACACACACAGGAGAGAACTCTCACTATTGAACACTGTCATCATACATCA 405
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 CTAGTCAGCGGGACACACAGCTATGATGCCCCGGTGAGCGAGTGGAGACACACCCC 405
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 406 CATTTCTCCACCAAGAACCAATGGACTATGATATTTGCCCTTTTGAAGATGGGTGGAGCC 465
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 406 CTGTACAGGGCAGCGCCCTCCA---CGCGCTGACGTGGCCCTGTGGAGCTGGAGGCACCA 462
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 466 TTCCAATTTGGGCACATTTGTGGGGCCCATATGTCCTCCAGAGCTGCGGGAGCAATTTGAG 525
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 GTCCCTTCACCAATTTACATCTCTCCCGGTGCTGCTGACCCCTGACCCCTCGGTGATCTTTGAG 522
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 526 GCTGGTTTTATTGTACAACATGCAGGCTGGGGCGGCTTAACCTGAAGGTGGCGTCTCTCA 585
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 523 ACGGGCATGAACGTGCTGGGTCACTGGCTGGGGCAGCCCCCAGAGGAGAACCTCTCTGCC 582
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 586 CAA 588
   ||
Db 583 GAA 585
   ||

RESULT 7
US-08-681-151-2
; Sequence 2, Application US/08681151
; Patent No. 5869637
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Braxton, Scott Michael
; APPLICANT: Goli, Surya
; TITLE OF INVENTION: A NOVEL HUMAN KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
```

```
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,151
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0074US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1739 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: HEARNOT01
; CLONE: 307474
; US-08-681-151-2

Query Match          5.6%; Score 51.8; DB 2; Length 1739;
Best Local Similarity 49.2%; Pred. No. 4.1e-07;
Matches 162; Conservative 1; Mismatches 163; Indels 3; Gaps 1;

QY 235 CATATTTTGGAGGAGCATCGTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATT 294
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1019 CATATTTGGGTGTTGTCCTCATTTGCCAAGAGTGGGTCTTGACAGTTGCCCATGCTTC 1078
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 295 GCAAAACAGAAACATTTGCTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAG 354
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1079 GAGGGGAGAGAGATGCTGCGAGTTTGGAAAGTGGTCTTGGGCATCAACAATCTAGACCAT 1138
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 355 ACAGACCCAGGAGAGCAAACTCTCACTATTTGAAACTGTCTATCATATACATCAATTTCTCC 414
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1139 CCATCAGTGTTCATGCAGACACGCTTTGTGAAGACCATCATCTGTCATCCCCG---CTAC 1195
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 415 ACCAAGAAACCAATGACATATGATATGCCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTT 474
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1196 AGTCGAGCAGTGGTGGACTATGACATCAGATCGTTTGAGCTGAGTGAAGACATCAGTGAG 1255
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 475 GGCCACTTTTGGGGGCCCATATGCTCTCCAGAGCTCGGGGAGCAATTTGAGGCTGGTTTT 534
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1256 ACTGGCTAGCTCGGGCTGTCTGCTTGCCCAACCCGAGCAGTGGCTAGAGCCTGACACG 1315
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 535 ATTTGTACAACATGCAGGCTGGGGCCGCTT 563
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1316 TACTGTCTATCATCAGGCTGGGGCCACAT 1344
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-08-807-151-2
; Sequence 2, Application US/08807151
; Patent No. 6043033
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL HUMAN PROSTATE-ASSOCIATED
; TITLE OF INVENTION: PROTEASE
```

Query Match 5.6%; Score 51.4; DB 3; Length 1077;
Best Local Similarity 49.0%; Pred. No. 4.2e-07;

	Matches	224; Conservative	0; Mismatches	226; Indels	7; Gaps	3;
QY	160	AGTCGCATTTCTGGAGGAAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTG	219			
Db	233	AGCAGGATCGTGGGCGGAGAGCGCGCTCCCGGGGCGTGGCCCTGGCAGGTGAGCCTG	292			
QY	220	AAACAAAGGACGAGGACATATTGTGGAGGAGGATCGTCTCACACACAGTGGGTGATCAGC	279			
Db	293	CAGTCCAGAACGTCCACGTGTGGGGAGGTCCCATCATCACCCCGAGTGGATCGTGACA	352			
QY	280	CGGCTCACGTGCATTTGCCAAACAGAAACATTTGTGCTA--CTTTGAAATGTTTACTGCTGGAG	337			
Db	353	CGCGCCACTCGCTGGGAAAAACCTCTTAAACAATCCATGGCATTTGCGGCGATTTCGCGGG	412			
QY	338	AGTATGACTTAAGCCACACAGACCCAGGAGGACAACTCTCAGTATTGAAACTGTCATCA	397			
Db	413	ATTTTGAGACAATCTTTTCATGTTCTATGGAGCCGGATACCAAGTA--GAAAAGAGTGATTT	470			
QY	398	TACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTTGAAGATGG	457			
Db	471	CTCATCCAAATTATGACTTCCAAGA--CCAGAACAAATGACATTTGGCTGATGAAGCTGC	527			
QY	458	CTGGAGCGCTTCCAAATTTGGGCACTTTGTGGGGCCCAATGATCTTCCAGAGCTCGGGGAGC	517			
Db	528	AGAACGCTCTGACTTTTCAACGACCTTAGTGAACACAGTGTGCTGCCCAACCCAGGCATGA	587			
QY	518	AATTGTAGGCTGGTTTTATTGTACAACTGCAGGCTGGGGCCGCTTAACCTAAGGTGGCG	577			
Db	588	TGCTGCAGCCACAGACACTCTGCTGGATTTCCGGGTGGGGGCCACCGAGGAGAAAGGA	647			
QY	578	TCCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCGCTAT	614			
Db	648	AGAGCTCAGAAAGTCTGAAACGCTCCCAAGGTGCTTCT	684			

[illegible]

```
QY 458 CTGAGCCTTCAATTTGGCCACTTTGTGGGGCCCAATGCTCTTCCAGAGCTGCGGGAGC 517
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 528 AGAAGCCTCTGACTCTCAACGAGCTAGTGAACACAGTGTCTGCCCAACCCAGGCATGA 587
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 518 AATTGAGGCTGTTTTATTGTACAACTCCAGGCTGGGGCCCTTAACTGAAGGTGGCG 577
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 588 TGTCTGAGCCAGAACAGCTCTGCTGGATTTCGGGTGGGGGCCACCCGAGGAGAAAGGGA 647
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 578 TCCTCTCACAAAGTTCGAGGAAGTGAATCTGCCTAT 614
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 648 AGACCTCAGAGTGTGACGAGCTGCCAAGGTGCTTCT 684
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-09-342-749-1
; Sequence 1, Application US/09342749
; Patent No. 6166194
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/342,749
; EARLIER FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 60/091,044
; EARLIER FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1476)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (724)
; OTHER INFORMATION: Listed as T in GenBank Accession No. U75329
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (985)
; OTHER INFORMATION: Listed as C in GenBank Accession No. 6166194 U75329
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (1347)
; OTHER INFORMATION: Listed as C in GenBank Accession No. 6166194 U75329
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (1466)
; OTHER INFORMATION: Listed as A in GenBank Accession No. 6166194 U75329
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (1471)
; OTHER INFORMATION: Listed as A in GenBank Accession No. 6166194 U75329.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (478)
; OTHER INFORMATION: This base can be G or A with G being the more
; OTHER INFORMATION: common allele. The codon will change from Val to
; FEATURE:
; NAME/KEY: allele
; LOCATION: (777)
; OTHER INFORMATION: This base can be C or T with C being the more
; OTHER INFORMATION: common allele. The codon is unaffected with both
; FEATURE:
; NAME/KEY: allele
; LOCATION: (768)
; OTHER INFORMATION: This base can be C or T with C being the more
```

```
; OTHER INFORMATION: common allele. This is a silent polymorphism.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (834)
; OTHER INFORMATION: This base can be C or T with C being the more
; OTHER INFORMATION: common allele. This is a silent polymorphism.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (625)
; OTHER INFORMATION: This base can be T or A with T being the more
; OTHER INFORMATION: common allele. The codon will change from Phe to
; OTHER INFORMATION: Ile
US-09-342-749-1

Query Match          5.6%; Score 51.4; DB 4; Length 1479;
Best Local Similarity 49.0%; Pred. No. 5.1e-07;
Matches 224; Conservative 0; Mismatches 226; Indels 7; Gaps 3;

QY 160 AGTCGCATTCTTTGGAGGAAGCCAAAGTGGAGAGCGCTTCTATCCCTGGCAGGTATCTCTG 219
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 760 AGCAGGATCGTGGCGCGGAGAGCGCGCTCCCGGGGGCCCTGGCCCTGGCAGGTACGCCCTG 819
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 220 AAACAAAGGCGAAGCATATTGTGTGAGGAAGCATCTCTACCAAGTGGGTGATCAGC 279
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 820 CAGTCCAGAACGTCACGCTGCGGAGGCTCCATCATCACCCCGAGTGGATCGTGACA 879
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 280 GGGGCTCACTGCATTCGAACAGAAACATTTGTCTTA--CTTTGAATGTTACTGCTGGAG 337
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 880 GCGGCCACTGCGTGGAAACCTCTTAACAATCCATGCGCATTTGGACGGCATTTGCGGGG 939
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 338 AGTATGACTTAAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATTGAACTGTGCATCA 397
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 940 ATTTGAGACAATCTTTTCATGTTCTATGGAGCGGATACCAAGTA--GAAAGATGATTT 997
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 398 TACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATGCCCCTTTTGAAGATGG 457
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 998 CTCATCCAAATATTGACTTCCAAGA---CCAAGAAACAATGACATTTGGCTGATGAAGCTGC 1054
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 458 CTGGAGCCTTCCAATTTGGCCCACTTTGTGGGGCCCAATATGCTCTCCAGAGCTGCGGGAGC 517
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1055 AGAAGCCTCTGACTTTCACAGACCTAGTGAACCACTGTGTGCCCCAACCCAGGCATGA 1114
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 518 AATTGAGGCTGTTTTATTGTACAACTGCAGGCTGGGGCCCTTAACTGAAGGTGGCG 577
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1115 TGTCTGAGCCAGAACAGCTCTGCTGGATTTCGGGTGGGGGCCACCCGAGGAGAAAGGGA 1174
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 578 TCCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCCTAT 614
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1175 AGACCTCAGAAAGTGTCTGAACGCTGCCAAGGTGCTTCT 1211
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-09-691-840-1
; Sequence 1, Application US/09691840
; Patent No. 6444419
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/691,840
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US/09/342,749
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/091,044
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1479
; TYPE: DNA
```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ..(1476)
NAME/KEY: conflict
LOCATION: (724)
OTHER INFORMATION: Listed as T in GenBank Accession NO. U75329
NAME/KEY: conflict
LOCATION: (985)
OTHER INFORMATION: Listed as C in GenBank Accession No. 6444419 U75329
NAME/KEY: conflict
LOCATION: (1347)
OTHER INFORMATION: Listed as C in GenBank Accession No. 6444419 U75329
NAME/KEY: conflict
LOCATION: (1466)
OTHER INFORMATION: Listed as A in GenBank Accession No. 6444419 U75329
NAME/KEY: conflict
LOCATION: (1471)
OTHER INFORMATION: Listed as A in GenBank Accession No. 6444419 U75329.
NAME/KEY: allele
LOCATION: (478)
OTHER INFORMATION: This base can be G or A with G being the more
OTHER INFORMATION: common allele. The codon will change from Val to
OTHER INFORMATION: Met.
NAME/KEY: allele
LOCATION: (777)
OTHER INFORMATION: This base can be C or T with C being the more
OTHER INFORMATION: common allele. The codon is unaffected with both
OTHER INFORMATION: alleles encoding Gly.
NAME/KEY: allele
LOCATION: (768)
OTHER INFORMATION: This base can be C or T with C being the more
OTHER INFORMATION: common allele. This is a silent polymorphism.
NAME/KEY: allele
LOCATION: (834)
OTHER INFORMATION: This base can be C or T with C being the more
OTHER INFORMATION: common allele. This is a silent polymorphism.
NAME/KEY: allele
LOCATION: (625)
OTHER INFORMATION: This base can be T or A with T being the more
OTHER INFORMATION: common allele. The codon will change from Phe to Ile
US-09-691-840-1

Query Match 5.6%; Score 51.4; DB 4; Length 1479;
Best Local Similarity 49.0%; Pred. No. 5.1e-07;
Matches 224; Conservative 0; Mismatches 226; Indels 7; Gaps 3;
QY 160 AGTCGCAATCTTGGAGGAGCCAAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTG 219
DB 760 AGCAGGATCGTGGCGGCGAGAGCGCGCTCCCGGGGCGCTGGCCCTGGCAGGTACGCCCTG 819
QY 220 AAACAAGGCAGNAGCATATTTGTGAGGAAGCATGCTCTCACCAGTGGGTGATCAGC 279
DB 820 CAGCTCCAGAGCTCCACGTGTGCGGAGGCTCCATCATCACCCTCCGAGTGGATGTGACA 879
QY 280 GGGCTCACTGCATTCACAAACAAACATTTGTCTTA--CTTTGAATGTACTGTGGAG 337
DB 880 GCGGCCACTCGCTGGAACAACTCTTAACAATCCATGCGATGGACGCAATTTGGCGGG 939
QY 338 AGTATGACTTTAAGCCAGACAGACCCAGGAGCAAACTCTCACTATTTGAACCTGATCA 397
DB 940 ATTTTGAGACAATCTTTTCACTGTCTATGAGCGCGGATACCAAGTA--GAAAGAATGATTT 997
QY 398 TACATCCACATTTCTCCACAGAACCAATGAGTATGATATGTCCTTTTGAAGATGG 457
DB 998 CTCATCAAAATATGACTCTCAAGA---CCAAGAACAATGACATTCGCTGATGAAGCTGC 1054
QY 458 CTGGAGCCTTCCAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGAGCTCGCGGAGC 517
DB 1055 AGAAGCCTCTGACTTTCAACGACCTAGTGAACCAAGTGTGTCTGCCCAACCCAGGCATGA 1114
QY 518 AATTTGAGGCTGTTTTATTTGTATCAACTGCAAGCTGCGGCGCGCTTAACTGAAGTGGCG 577

DB 1115 TGCTGCAGCCAGAACAGCTCTCTGGATTTCGGGGTGGGGCCACCGAGGAGAAAGGA 1174
QY 578 TCCTCTCACAAGCTTGCAGGAAGTGAATCTGCCTAT 614
DB 1175 AGACCTCAGAAGTGTGAACGCTGCCAAGTGTCTT 1211
RESULT 12
US-09-088-651-1
; Sequence 1, Application US/09088651
; Patent No. 6165771
; GENERAL INFORMATION:
; APPLICANT: BURGESS, NICOLA A.
; APPLICANT: CLINKENBEARD, HELEN E.
; APPLICANT: SOUTHAN, CHRISTOPHER D.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,651
; FILING DATE: JUNE 1, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB9712088.5
; FILING DATE: 10-JUNE-1997
; APPLICATION NUMBER: EP 97308295.1
; FILING DATE: 17-OCT-1997
; APPLICATION NUMBER: GB 9803650.2
; FILING DATE: 20-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F.
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH30358
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-088-651-1

Query Match 5.5%; Score 51; DB 4; Length 1109;
Best Local Similarity 46.4%; Pred. No. 5.8e-07;
Matches 202; Conservative 0; Mismatches 230; Indels 3; Gaps 1;
QY 154 ATTTTCAGTCGCAATCTTGGAGGAAGCCAAAGTGGAGAGGTTCTCTATCCCTGGCAGGTA 213
DB 287 ATGCTGAACCGAATGGTGGCGGGCAGGACACGACGAGGAGGCGGAGTGGCCCTGGCAAGTC 346
QY 214 TCTCTGAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 273
DB 347 AGCATCCACCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 406
QY 274 ATCAGCGGCGCTCACTGTCATTCGAAACAGAAACATTTGTCTACTTTTGAATGTTACTGCT 333
DB 407 CTGACGGCTGCGCACTGCTTCCGCAACACCTCTGAGACGCTCCCTGGACGAGGCTCTCTGCTG 466

Best Local Similarity	52.9%;	Pred. No.	9.8e-07;
Matches	165;	Conservative	0;
		Mismatches	132;
		Indels	15;
		Gaps	2;
<hr/>			
QY	192	GGGTTCCTATCCCTGGCAGGTTACTCTGAACAAGAAGCAGAAGCATATTTTGTGGGGAAG	251
Db	832	GGATTCTTGCCCTTGGCAGGTGCAGCATCCAGTAGCAACAACAGCACGCTCTGTGGAGGAG	891
<hr/>			
QY	252	CATCGTCTCACCAGATGGGTGATCAGCGGGCTCACTGCATTCGCAACACAGAACAATTGT	311
Db	892	CATCCTGGACCCCCACATGGGTCTTCAGCGCAGCCATGCTTCAGGAACAATACCGATGT	951
<hr/>			
QY	312	GTCTACTTTGAATGTTACTGTGTGAGAGTAGTACTTAAGCCAGACAGACCCAGGAGAGCA	371
Db	952	GTTCAACTGGAAGGTGGCGCAGGCTCAGACAAACTGGCA-----GCTTCCC	999
<hr/>			
QY	372	AAC TCTC A C T A T T G A A A C T G T C A T C A T A C A T T T C T C C A C C A A G A A C C A A T G G A	431
Db	1000	A T C C T G G C T G T G C C C A A G A T C A T C A T T G A A --- T T C A A C C C C A T G T A C C C C A A G A	1056
<hr/>			
QY	432	CTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTGTGGGGCC	491
Db	1057	CAATGACATGCCCTCATGAAGCTGCAGTCTCCCACTCACTTTCTCAGGCACAGTCAGGCC	1116
<hr/>			
QY	492	CATA TGCTTCC	503
Db	1117	CATCTGCTGCC	1128

Search completed: March 25, 2003, 06:31:46
Job time : 70 secs

RESULT 15
 US-09-656-002-1
 ; Sequence 1, Application US/09656002
 ; Patent No. 6455668
 ; GENERAL INFORMATION:
 ; APPLICANT: Mack, David
 ; APPLICANT: Gish, Kurt
 ; APPLICANT: Wilson, Keith
 ; TITLE OF INVENTION: NOVEL METHODS OF SCREENING FOR COLORECTAL CANCER, COMPOSITIONS, AND
 ; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
 ; FILE REFERENCE: A-69108/DJB/JJD/AWS
 ; CURRENT APPLICATION NUMBER: US/09/656.002
 ; CURRENT FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: US 09/525,993
 ; PRIOR FILING DATE: 2000-03-15
 ; PRIOR APPLICATION NUMBER: US 09/493,444
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: PCT/US 00/07044
 ; PRIOR FILING DATE: 2000-03-15
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2079
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-656-002-1

Query Match.	5.5%	Score 50.8;	DB 4;	Length 2079;
Best Local Similarity	52.9%	Pred. No. 9.9e-07;		
Matches 165;	Conservative	0;	Mismatches 132;	Indels 15; Gaps 2;

QY	192	GGGTTCTATCCCTGGCAGGTATCTCTGAACAAGGCAGAAAGCATATTTCTGGAGGAAG	251
Db	847	GGATTTCTGGCCTTGGCAGGTCAGCATCCAGTACGACAAACACGCGTCTCTGGAGGGAG	906
QY	252	CATCGTCTCACCAGTGGGTGATCAGCGGGCTCACTGCATTCGAACAGAAACATGT	311
Db	907	CATCTGGACCCCACTGGGTCTCAGCGGAGCCCACTGCTTCAGGAACAACATCCGATGT	966
QY	312	GTCTACTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGACCAGGAGAGCA	371
Db	967	GTCAACTGGAAGGTGGCGGAGGCTCAGACAAACTGGGCA-----GCTTCCC	1014
QY	372	AACTCTCACTATTGAAACTGTCACTCATCATCCACATTTTCTCCACGAAGAACAATGGA	431

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 25, 2003, 01:26:19 ; Search time 3423 Seconds
(without alignments)
2601.653 Million cell updates/sec
Title: US-09-735-713a-2
Perfect score: 1636
Sequence: 1 MSLKMLISRNKLLILLGIVF.....IFTDISKVLWTHIQTGN 306

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09735713/runat_18032003_124141_21590/app_query.fasta_1.455
-DB-GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713 @CNG1.1.2496 @runat_18032003_124141_21590 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rtd.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1632	99.8	1262	6	AX342644	Sequence
2	1626	99.4	921	6	AX180293	Sequence
3	1626	99.4	1568	6	AX180299	Sequence
4	1608	98.3	909	6	AX180295	Sequence
5	1372.5	83.9	1671	6	AX360091	Sequence
6	898	54.9	495	6	AX180297	Sequence
7	831	50.8	164732	9	AC104237	Homo sapi
8	831	50.8	180707	2	AC012228	Homo sapi
9	688.5	42.1	3689	5	XL081291	Xenopus lae
10	657	40.2	3028	5	AB070367	Bufo japo
11	626	38.3	4828	5	XL081290	Xenopus lae
12	576	35.2	2627	6	AX480940	Sequence
13	529.5	32.4	169388	2	AC129620	Rattus no
14	496	30.3	942	3	SSU79521	Scolopendra
15	489.5	29.9	2409	6	AX360098	Sequence
16	486.5	29.7	2672	6	AX207905	Sequence
17	486.5	29.7	3104	6	AX207903	Sequence
18	486	29.7	2594	3	AY119618	Drosophila
19	486	29.7	3749	3	DROSTUBBLE	Drosophila
20	480.5	29.4	3183	10	BC029645	Mus muscu
21	476.5	29.1	708	6	AX207901	Sequence
22	475.5	29.1	1595	10	AB052292	Mus muscu
23	475.5	29.1	2231	10	BC019485	Mus muscu
24	469.5	28.7	2025	10	AF356627	Mus muscu
25	467	28.5	1359	6	AX354827	Sequence
26	467	28.5	1656	6	AX207907	Sequence
27	467	28.5	2412	6	AX375996	Sequence
28	467	28.5	2948	6	AX354825	Sequence
29	466	28.5	2267	6	AX335042	Sequence
30	466	28.5	2267	6	AX409602	Sequence
31	466	28.5	2267	9	HUMPPKKA	Nucleotide
32	464	28.4	2753	10	MUSPKA	Mouse plas
33	462.5	28.3	1801	3	AF357226	Panulirus
34	462	28.2	2752	10	MMU300738	Mus muscu
35	458	28.0	2398	10	BC026555	Mus muscu
36	456.5	27.9	1365	6	AX360088	Sequence
37	456.5	27.9	1365	6	AX392866	Sequence
38	456.5	27.9	2135	6	AX277411	Sequence
39	456.5	27.9	2135	9	AB038157	Homo sapi
40	456.5	27.9	2468	9	AB038157	Homo sapi
41	456.5	27.9	2554	9	AB038158	Homo sapi
42	456.5	27.9	2951	6	AX392864	Sequence
43	454.5	27.8	3387	6	AX360082	Sequence
44	450	27.5	1743	10	BC003851	Mus muscu
45	450	27.5	1753	10	AF378085	Mus muscu

ALIGNMENTS

AX342644 AX342644 1262 bp DNA linear PAT 12-JAN-2002
LOCUS Sequence 41 from Patent WO0198468.
ACCESSION AX342644
VERSION AX342644.1 GI:18152041
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Elliott,V.S., Gandhi,A.R., Lal,P., Au-Young,J.,
Tribouley,C.M., Deleage,A.M., Baughn,M.R., Nguyen,D.B., Lee,E.A.,
Hafalia,A., Khan,F.A., Walia,N.K., Yao,M.G., Lu,D.A., Patterson,C.,
Tang,Y.T., Walsh,R.T., Azimzai,Y., Ramkumar,J., Xu,Y. and Reddy,R.
JOURNAL Patent: WO 0198468-A 41 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 1262
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 433459Cbl"
BASE COUNT 354 a 265 c 306 g 337 t
ORIGIN
Alignment Scores:
Pred. No.: 1.15e-144 Length: 1262
Score: 1632.00 Matches: 305
Percent Similarity: 99.67% Conservative: 0
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 99.76% Indels: 0
DB: 6 Gaps: 0
US-09-735-713A-2 (1-306) x AX342644 (1-1262)
Qy 1 MetSerLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuLeuLeuValPhe 20
Db 165 ATGAGTCTCAAAATGCTTATAGCAGGAACAAGCTGATTACTAGGAATAGCTTT 224
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyClnSer 40
Db 225 TTGNACGAGGTAAATCTGCAACTCTTTTCGCTCCCAAGCTCCAGTGTGGCAGAGT 284
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
Db 285 CTGGTTAAGGTACAGCCTTGGAATATTATTAACATTTTCAGTCGCATCTTTGGAGGAGC 344
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 345 CAAGTGGAAGAAGGTTCTATCCCTGGCAGGTATCTCTGAACAACAAAGCAGACATATT 404
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 405 TGTGGAGAACATCGTCTCACACAGTGGGTGATCAGCGGGCTCACTGCTATGGCAAC 464
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Db 465 AGAACATTTGTCTACTTTTCAATGTTACTCTGCGAGAGTATGACTTAAGCCAGACAGAC 524
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 525 CCAGAGAGCAACTCTCACATTATGAAGTGTGATCATCATCATCATCATCATCATCAT 584
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 585 AAACCAATGGACTATGATATGCCCTTTTGAAGTGGCTGGAGCCTTCCAATTTGGCCAC 644
Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 645 TTTGTGGGGCCCATATGCTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTTTATTGT 704
Qy 181 ThrThrAlaGlyTyrGlyArgLeuThrClnGlyGlyValLeuSerGlnValLeuGlnGlu 200

Db 705 ACAACTGCAGCTGGGGCCGCTTAACCTGAAGGTGGCTCTCTCAAAAGTCTTGCAGGAA 764
Qy 201 ValAsnLeuProIleLeuLeuThrTrpGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 765 GTGAATCTGCCTATTTTACCTGGGAAGAGTGTGGGAGCTCTGTTAACACTAAAGAGG 824
Qy 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 825 CCCATCAGTGGGAAGACCTTTCTTTCAGAGGTTTCTGATGGAGGAGAGACGATGT 884
Qy 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Db 885 CAGGAGATTCAGGAGGTCTACTCATGTGCGGAATAAGAAAGGGCCCTGGACTCTGGCT 944
Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTyrPheAsnValArgLysSer 280
Db 945 GGTGTGACTTCTGGGGTTTGGGTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGT 1004
Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Db 1005 GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAGTGTCTTCTGGATCCAGAA 1064
Qy 301 HisIleGlnThrGlyAsn 306
Db 1065 CACATCCAAACTGGTAAC 1082
RESULT 2
AX180293
LOCUS AX180293 921 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0146407.
ACCESSION AX180293
VERSION AX180293.1 GI:15132262
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Polynucleotides encoding human,protease homologs
JOURNAL Patent: WO 0146407-A 1 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1. 921
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 242 a 195 c 240 g 242 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 2.9e-144 Length: 921
Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 6 Gaps: 0
US-09-735-713A-2 (1-306) x AX180293 (1-921)
Qy 1 MetSerLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuLeuValPhe 20
Db 1 ATGAGTCTCAAAATGCTTATAGCAGGAACAAGCTGATTACTAGGAATAGCTTT 60
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyClnSer 40
Db 61 TTTGAACRAGGTAAATCTGCARCTCTTTCGCTCCCAAGCTCCAGTGTGGCAGAGT 120
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
Db 121 CTGGTTAAGGTACAGCCTTGAATATTATTAACATTTTCAGTCGCATCTTTGGAGGAGC 180
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80

Db	181	CAAGTGGAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGAGAGCATATT	240
QY	81	CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn	100
Db	241	TGTGGAGGAAGCATCGTCTCACACAGTGGGTGATCAGCGGCTGCATGCAATGCAAC	300
QY	101	ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp	120
Db	301	AGAAACATTGTGTCTACTTTGAATGTTACTGTGGAGAGTATGACTTAAAGCCAGACAGAC	360
QY	121	ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys	140
Db	361	CCAGGAGAGCAACTCTCACTATTGAACCTGTCATCATCATCATCTTCCACCAAG	420
QY	141	LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis	160
Db	421	AAACCAATGGACATGATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCAATTTGGCCAC	480
QY	161	PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys	180
Db	481	TTTGTGGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTGT	540
QY	181	ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu	200
Db	541	ACAACCTGCAGGCTGGGGCCGCTTAACCTGAAGGTGGCGTCTCTCACAAAGTCTTGCAGGA	600
QY	201	ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg	220
Db	601	GTGAATCTGCCATATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG	660
QY	221	ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys	240
Db	661	CCCATCAGTGGGAAGACCTTCTTTCACAGAGTTTTCCTGATGGAGGAGAGACGATGT	720
QY	241	GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaIleThrPheLeuAla	260
Db	721	CAGGAGATTCAGGAGGTTCACTCATGTGCCGAATTAAGAAAGGGCCCTGGACTCTGGCT	780
QY	261	GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer	280
Db	781	GGTGTCACTCTCCGGGTTTGGCTGTGTGGAGGCTGGAGAAACAATGTGAGAGAAAGT	840
QY	281	AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu	300
Db	841	GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTGCTTTCCTGGATCCAGGA	900
QY	301	HisIleGlnThrGlyAsn	306
Db	901	CACATCCAAACTGGTAAC	918
RESULT 3			
AX180299			
LOCUS	AX180299 1568 bp DNA linear PAT 06-AUG-2001		
DEFINITION	Sequence 7 from Patent WO0146407.		
ACCESSION	AX180299		
VERSION	AX180299.1 GI:15132265		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1. (bases 1 to 1568)		
AUTHORS	Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and Sands,A.I.		
TITLE	Polynucleotides encoding human protease homologs		
JOURNAL	Patent: WO 0146407-A 7 28-JUN-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..1568		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	452 a	338 c	423 t 2 others

ORIGIN			
Alignment Scores: 5,47e-144 Length: 1568			
Pred. No.: 1626.00 Matches: 304			
Score: 1626.00			
Percent Similarity: 99.35% Conservative: 0			
Best Local Similarity: 99.35% Mismatches: 2			
Query Match: 99.39% Indels: 0			
DB: 6 Gaps: 0			
US-09-735-713A-2 (1-306) x AX180299 (1-1568)			
QY	1	MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe	20
-Db	245	ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT	304
QY	21	PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer	40
Db	305	TTTGAACRAGGTAAATCTGCARCTCTTGGCTCCCCAAAGCTCCCAAGTTGTGGGAGAGT	364
QY	41	LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer	60
Db	365	CTGGTTAAGGTACAGCCTTGGAAATTATTTAACATTTTCAGTCGCATTTCTGGAGGAAGC	424
QY	61	GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle	80
Db	425	CAAGTGGAGAAGGGTTCCTATCCCTGCGCAGGTATCTCTGAACAAAGGCGAAGACATATT	484
QY	81	CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn	100
Db	485	TGTGGAGGAAGCATCGTCTCACACAGTGGGTGATCAGCGGCTCCTCATTCATTTGCAAC	544
QY	101	ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp	120
Db	545	AGAAACATTGTGTCTACTTTGAATGTACTGCTGGAGAGTATGACITTAACCCAGACAGAC	604
QY	121	ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys	140
Db	605	CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATCATCATCATTTCTCCACCAAG	664
QY	141	LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis	160
Db	665	AAACCAATGGACATGATATFTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC	724
QY	161	PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys	180
Db	725	TTTGTGGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTGT	784
QY	181	ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu	200
Db	785	ACAACTGCAGGCTGGGGCCGCTTAACTGAAGGTGGCGTCTCTCACAAAGTCTTGCAGGAA	844
QY	201	ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg	220
Db	845	GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTTAACACTAAAGAGG	904
QY	221	ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys	240
Db	905	CCCATCAGTGGGAAGACCTTCTTGTGCACAGGTTTCTCTGATGGAGGAGAGACGATGT	964
QY	241	GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaIleThrPheLeuAla	260
Db	965	CAGGAGATTCCAGAGGTTTCACTCATGTGCGCGGAATAAGAAAGGGGCTGAGCTCTGGCT	1024
QY	261	GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer	280
Db	1025	GGTGTCACTTCTCCGGGTTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGT	1084
QY	281	AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu	300
Db	1085	GATCAAGGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGCTTTCCTGGATCCAGGA	1144
QY	301	HisIleGlnThrGlyAsn	306

```

Db 1145 CACATCCAAACTGGTAAC 1162
RESULT 4
AX180295
LOCUS AX180295 909 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 3 from Patent WO0146407.
ACCESSION AX180295
VERSION AX180295.1 GI:15132263
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS Waite,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Polynucleotides encoding human protease homologs
JOURNAL Patent: WO 0146407-A 3 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..909
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 237 a 193 c 238 g 239 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.42e-142 Length: 909
Score: 1608.00 Matches: 300
Percent Similarity: 99.34% Conservative: 0
Best Local Similarity: 99.34% Mismatches: 2
Query Match: 98.29% Indels: 0
DB: 6 Gaps: 0
US-09-735-713a-2 (1-306) x AX180295 (1-909)
QY 5 MetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPheGluArgGly 24
Db 1 ATGCCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTTTGAACRAGGT 60
QY 25 LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal 44
Db 61 AAATCTGCARCTCTTCCTGCCAAAGCTCCCAAGCTCCAGTGTGGCAGAGTCTGGTTAAGTA 120
QY 45 GlnProTyrAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLys 64
Db 121 CAGCCTTGAATATTTTAAACATTTTCAGTCGATCTTGGAGGAAGCAAGTGGAGAG 180
QY 65 GlySerTyrProTyrGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySer 84
Db 181 GGTTCCTATCCCTGGCAGGATCTCTGAACAAGCAAGCAGACATATTTGTGGAGGAAGC 240
QY 85 IleValSerProGlnTyrValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal 104
Db 241 ATCGTCTCACACAGTGGGTGATCAGCGGCTCACTGCATTCGCAACAGAAACATTTGTG 300
QY 105 SerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGln 124
Db 301 TCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCA 360
QY 125 ThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysProMetAsp 144
Db 361 ACTCTCACTATTTGAACCTGTCATCATCATCATCATTTCTCCACCAAGAAACCAATGGAC 420
QY 145 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 164
Db 421 TATGATATGGCCCTTTTGAAGATGCTGGAGCTTCCCAATTTGGCCACTTTGTGGGGCCC 480
QY 165 IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 184
Db 481 ATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGTTTATTTTGTACAACTGCAGGC 540

```

```

QY 185 TrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGluValAlaAsnLeuPro 204
Db 541 TGGGGCCGCTTAACCTGAAGGTGGCGCTCTCTCACAGTCTTCGAGGAAGTGAATCTGCCT 600
QY 205 IleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIleSerGly 224
Db 601 ATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGG 660
QY 225 LysThrPheLeuCysThrClyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSer 244
Db 661 AAGACCTTTCTTTTGACAGGTTTTCTGATGAGGAGAGAGCATGTCTCAGGAGATTCA 720
QY 245 GlyGlySerLeuMetCysArgAsnLysLysGlyAlaIleThrLeuAlaGlyValThrSer 264
Db 721 GGAGGTTCACTCATGTGCGGAATAAGAAAGGGGCTTGACTCTGGCTGGTGTGACTTCC 780
QY 265 TrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySer 284
Db 781 TGGGGTTTGGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGGAAAGTATCAAGGATCC 840
QY 285 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 304
Db 841 CCTGGGATCTTCACAGACATTAAGTAAAGTCTTCTGGATCCACAGACATCCAAACT 900
QY 305 GlyAsn 306
Db 901 GGTAAC 906
RESULT 5
AX360091
LOCUS AX360091 1671 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 47 from Patent WO0200860.
ACCESSION AX360091
VERSION AX360091.1 GI:18675717
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
Charydczak,G.
TITLE Novel proteases
JOURNAL Patent: WO 0200860-A 47 03-JAN-2002;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1671
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 451 a 369 c 404 g 447 t
ORIGIN
Alignment Scores:
Pred. No.: 4.93e-120 Length: 1671
Score: 1372.50 Matches: 266
Percent Similarity: 93.75% Conservative: 4
Best Local Similarity: 92.36% Mismatches: 6
Query Match: 83.89% Indels: 12
DB: 6 Gaps: 2
US-09-735-713a-2 (1-306) x AX360091 (1-1671)
QY 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20
Db 1 ATGAGTCTCAAAATGCTTTATAAGCAGGAAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
QY 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 61 TTTGAACGAGGTAATCTGCAACTCTTTTCGCTCCCAAGACTCCCAAGTGTGTGGCAGAGT 120
QY 41 LeuValLysValGlnProTyrAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
Db 121 CTGGTTAAGGTACAGCCTTGGAAATATTTTAACTTTTTCAGTCGCATTTCTTGGAGGAAGC 180

```

QY	61	GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle	80
Db	181	CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAGGAGAGCATATT	240
QY	81	CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn	100
Db	241	TGTGGAGGAGCATCGTCTCACCACAGTGGGTGATCAGCGGCTCAGTCATTGCAAAC	300
QY	101	ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp	120
Db	301	AGAAACATTGTGTCTACTTTGAATGTACTGCTGGAGAGTATGACATTAAGCCAGACAG	360
QY	121	ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys	140
Db	361	CCAGGAGAGCAAACTCTCACTATTGAACGTCTCATCATATCCACATTTCTCCACCAAG	420
QY	141	LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis	160
Db	421	AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC	480
QY	161	PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys	180
Db	481	TTTGTGGGGCCCATATGCTCTCCAGAGCTGCGGGAGCAATTTTGAGGCTGGTTTTATTGT	540
QY	181	ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu	200
Db	541	ACAACCTGCAGCTGGGGCCCTTAACCTGAAGTGGCGTCTCTCAACAAGTCTTGCAAGAA	600
QY	201	ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg	220
Db	601	GTCAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGCAGCTCTGTTAACACTTAAGAGG	660
QY	221	ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys	240
Db	661	CCCATCAGTGGGAAGACCTTCTTTGCACAGGTTTTCCTGATGGAGGAGAGCAGCATGT	720
QY	241	GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrp-ThrLeuAl	260
Db	721	CAGGAGATTCAGAGGTTCCATCTGTCGGGAATAAGAAAGGGCCCTGGGACTCTGGC	780
QY	260	aglyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSe	280
Db	781	TGGT-----CAATTTGGG-----AGGCTCAGGTGGGA	807
QY	280	rAspGlnGlySerProGlyIle	287
Db	808	GGATCGCTTGAGTCCAGGAGTT	829
RESULT 6			
LOCUS	AX180297	495 bp	DNA linear PAT 06-AUG-2001
DEFINITION	Sequence 5 from Patent WO0146407.		
ACCESSION	AX180297		
VERSION	AX180297.1	GI:15132264	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 495)		
TITLE	Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and		
JOURNAL	Sands,A.T.		
FEATURES	Polynucleotides encoding human protease homologs		
Source	Patent: WO 0146407-A 5 28-JUN-2001;		
BASE COUNT	Lexicon Genetics Incorporated (US)		
ORIGIN	Location/Qualifiers		
	1..495		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	117 a 101 c 147 g 130 t		

Alignment Scores:	6.87e-76	Length:	495
Pred. No.:	898.00	Matches:	164
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	54.89%	Indels:	0
DB:	6	Gaps:	0
US-09-735-713A-2 (1-306) x AX180297 (1-495)			
QY	143	MetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal	162
Db	1	ATGGAATATGATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTGTG	60
QY	163	GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr	182
Db	61	GGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTTGAGGCTGGTTTTATTGTACAACT	120
QY	183	AlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnValAsn	202
Db	121	GCAGGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCAACAAGTCTTGCAAGAGTGAAT	180
QY	203	LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle	222
Db	181	CTGCCATTTTGAACCTGGGAAGAGTGTGTGCAGCTCTGTTAACACTTAAAGAGGCCCATC	240
QY	223	SerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgaspAlaCysGlnGly	242
Db	241	AGTGGGAAGACCTTCTTTCACAGGTTTTCCTGATGGAGGAGAGCAGCATGTGTCAGGA	300
QY	243	AspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyVal	262
Db	301	GATTCAGGAGGTTCACTCATGTCGGGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG	360
QY	263	ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln	282
Db	361	ACTTCCTGGGCTTTGGCTGTGTCGAGGCTGGAGAAACAATGTGAGGAAAGTGTACAA	420
QY	283	GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle	302
Db	421	GGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGCTTCTTCTGGATCCAGAACATC	480
QY	303	GlnThrGlyAsn	306
Db	481	CAACCTGGTAAAC	492
RESULT 7			
LOCUS	AC104237	164732 bp	DNA linear PRI 30-APR-2002
DEFINITION	Homo sapiens chromosome 11, clone RP11-35J10, complete sequence.		
ACCESSION	AC104237		
VERSION	AC104237.2	GI:20128277	
KEYWORDS	HTG.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 164732)		
JOURNAL	Birren,B., Linton,L., Nusbaum,C. and Lander,E.		
REFERENCE	Homo sapiens chromosome 11, clone RP11-35J10		
AUTHORS	Unpublished		
	2 (bases 1 to 164732)		
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,		
	Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,		
	Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,		
	Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,		
	Cooke,P., DeAtrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,		
	Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,		
	Grinde,S., Gord,S., Goyette,W., Graham,L., Grand-Pierre,N.,		
	Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,		
	Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,		
	Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,		
	MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,		

McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (06-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164732)

REFERENCE AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhaltier,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., MacClean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164732)

REFERENCE AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhaltier,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., MacClean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (30-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 10, 2002 this sequence version replaced 17386405.

COMMENT

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21899
Center clone name: 35_J_10

FEATURES

source

Location/Qualifiers
1. .164732
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="L1"
/clones="RP11-35J10"
/clone_lib="RPC1-11 Human Male BAC"
5. .11

unsure

/note="<30 qual SNGL region"
126. .364
/rpt_family="L1MBC"
complement(388. .1255)
/rpt_family="L1P"
complement(1256. .1759)
/rpt_family="MLT2A2"
complement(1901. .2313)
/rpt_family="MLTIC"
complement(4166. .4174)
/note="<30 qual SNGL region"
4381. .4688
/rpt_family="AluSx"
6150. .6445
/rpt_family="AluSx"
6449. .6575
/rpt_family="L2"
7433. .7667
/rpt_family="MIR"
complement(8640. .9715)
/rpt_family="L1"
complement(9716. .10025)
/rpt_family="AluJo"
complement(10026. .10141)
/rpt_family="L1"
complement(10142. .10443)
/rpt_family="AluSq"
complement(10444. .11104)
/rpt_family="L1"
complement(11118. .11295)
/rpt_family="L1"
complement(11308. .17289)
/rpt_family="L1PA2"
complement(17291. .17784)
/rpt_family="L1"
complement(17801. .19730)
/rpt_family="L1PB3"
19747. .20463
/rpt_family="L1M4"
20464. .20747
/rpt_family="AluSx"
20748. .21462
/rpt_family="L1M4"
complement(21788. .22016)
/rpt_family="L1M4"
22017. .22111
/rpt_family="CT-rich"
22219. .22255
/rpt_family="(CA)n"
complement(24696. .24856)
/rpt_family="MIR"
complement(24993. .25122)
/rpt_family="L2"
25237. .25470
/rpt_family="L2"
complement(25810. .25919)
/rpt_family="MIR"
28709. .29154

repeat_region

repeat_region	/rpt_family="LRI16A1"	
repeat_region	29252..29284	
repeat_region	/rpt_family="(CAAAA)n"	
repeat_region	29612..30220	
repeat_region	/rpt_family="L1"	
repeat_region	30219..30485	
repeat_region	/rpt_family="L1"	
repeat_region	30486..30788	
repeat_region	/rpt_family="AlusX"	
repeat_region	complement(30809..31550)	
repeat_region	/rpt_family="L1P"	
repeat_region	31551..32627	
repeat_region	/rpt_family="L1PA3"	
repeat_region	complement(32836..33451)	
repeat_region	/rpt_family="L1P"	
repeat_region	complement(33489..33965)	
repeat_region	/rpt_family="L1"	
repeat_region	33977..34152	
repeat_region	/rpt_family="L1"	
repeat_region	complement(34157..34342)	
repeat_region	/rpt_family="MER46A"	
repeat_region	34343..35088	
repeat_region	/rpt_family="L1MDa"	
repeat_region	35089..35303	
repeat_region	/rpt_family="L2"	
repeat_region	35968..36617	
repeat_region	/rpt_family="L1PA13"	
unsure	36437..36505	
	/note="single clone coverage"	
Alignment Scores:		
Pred. No.:	1,49e-66	Length: 164732
Score:	831.00	Matches: 205
Percent Similarity:	30.21%	Conservative: 1
Best Local Similarity:	30.06%	Mismatches: 1
Query Match:	50.79%	Indels: 475
DB:	9	Gaps: 3
US-09-735-713A-2 (1-306) x AC104237 (1-164732)		
Qy 100	AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThr	119
Db	49447 CATAGAAACATGGTGTCTACTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACA	49506
Qy 120	AspProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThr	139
Db	49507 GACCCAGGAGCAAACTCTCATTGAAACTGCTCATCATATACATCCACATTTCTCCACC	49566
Qy 140	LysLysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPhe---	158
Db	49567 AAGAACCAATGGACTATGATATTGCCCTTTTGAAGATGCTGGAGCCTTCCAATTTGGT	49626
Qy 158	-----	158
Db	49627 AACATTGGAGATACATACTGAATTGTCTACCCCATGTTCTCTATACCTCTGTGCCATA	49686
Qy 158	-----	158
Db	49687 GCACATTTCTGCAGTATACGTTCTGGGTCTGCGTCATGACATTTGAGATAGGAATTTAGGCTGGA	49746
Qy 158	-----	158
Db	49747 GAGTCCTACCTGAAACTCTTTTTCGTTGTTCTGGAGATAGCATTACATGGAGTCAGG	49806
Qy 158	-----	158
Db	49807 AGAAGGTTTCTTTAATTGACAGATAACCCCTGGAGGTGCCAGGACGAGCTATCACAACT	49866
Qy 158	-----	158
Db	49867 CTCATGAGGCTTTTAGGGGACCACCTGGATACCAAGTGGTGGCTTATAGTGTCTC	49926
Qy 159	-----GlyHisPheValGlyPro	164

Db 49927	TCCTCCATGAATGAGACAAAGCCCTACCTTGCTCTTTTCTTTCAGGCCACTTTGTGGGCC	49986
Qy 165	IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly	184
Db 49987	ATATGCTCTCCAGAGCTGCGGAGCAATTTGAGCTGGTTTTATTTGTACAACTGCAGGC	50046
Qy 185	TrpGlyArgLeuThrGlu-----	190
Db 50047	TGGGGCCGCTTAAGGTAAGAACTTGTGTCCCATTTCTCTTAAGCAGAGGTCAGAA	50106
Qy 190	-----	190
Db 50107	CTTTCTGGTGGCAATCAGAAATGGAATTTTAAATCCATTTTAAAAATCGATTGTGCA	50166
Qy 190	-----	190
Db 50167	TACATTTCTAAAGATAGGGAATGTACATTTGTAAGAAATTCAGTCTTGTCCAAGCGA	50226
Qy 190	-----	190
Db 50227	CACAAACAAATGATGCATCTTCAACTCAGAAAGGGTGAAGAGTTCCTCTGGGGCTT	50286
Qy 191	-----GlyGlyValIleuSerGlnValLeuGlnGluValIasnLeuPro	204
Db 50287	CCATCTCTCTGTTTTTCAGGTGGCGCTCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCCT	50346
Qy 205	IleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIleSerGly	224
Db 50347	ATTTTGACCTGGGAAGAGTGTGTGCAGCTCTGTTAACTAAAGAGGCCATCAGTGGG	50406
Qy 225	LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys-----	240
Db 50407	AAGACCTTTCTTGCACAGGTTTTCTCTGATGGAGGAGAGACGATGTACAGGTGGTGGG	50466
Qy 240	-----	240
Db 50467	AGCTGGGGGCTGTGAGGCCCCAGGCAAGGATTCCTTGGTTTGTCTTCTTCTAGTTTT	50526
Qy 240	-----	240
Db 50527	CTTAAATTTTCAGTGAAATTTAAATTAATTAAGAAAGCATATTGGAGGGGTAGAATAAGAA	50586
Qy 240	-----	240
Db 50587	TAGTCTTGAAGAAGCTATAAGGGTCAAAATAGTTAGTTAAATGTGTGAGAAAAATGGTAGA	50646
Qy 240	-----	240
Db 50647	TCCAAACCATTTCCCTTAAGAGGCTCACCTGTTATCAGGTGAGGGGCTCCAGGAAGACTGA	50706
Qy 240	-----	240
Db 50707	TTCAGGTAAGGCAGGAATCTGGAAGAAGGGTTAGGAGTACCAAGGAGGTTACATCAA	50766
Qy 240	-----	240
Db 50767	AATAAAGGACAGTGGAAACAATCTGTTTTAGCAGGAAGTGTGATTTAAAGAGGCTTT	50826
Qy 240	-----	240
Db 50827	CTGCTTTTCCAAAAGACCAATAAACAGAGATAATGCCAAGAAATGAACCTCGGGGTTCTCC	50886
Qy 240	-----	240
Db 50887	CATTCTGAAGGATGGCTGACTCTTCTTATTTAGAAATTTTCCAGTTTGTGGATGTTGGG	50946
Qy 240	-----	240
Db 50947	AACTTTACACTTTTCTCCCATTTCTTCTACTAAACCAAGGCCGTAAGTCAGTCTCATTTGCT	51006
Qy 240	-----	240

```

Db 51007 GGCAAGACTTCTCTGAGTCTGCAGCATGAGAGAGAGAGAAAGAAAGTCTGCTGCTTT 51066
QY 240 ----- 240
Db 51067 GATGAAGCCATGATGAGAGAAAGGGAGGTGGGATTTGTTGGTTCACATGATAGTTGAG 51126
QY 240 ----- 240
Db 51127 GATCATATGGGAATTTGCTGGGATCTGTATACAGAGCCCTGAGTCTAAATATTAGA 51186
QY 240 ----- 240
Db 51187 ACTGAGAGATGTCATTGATCTCTCCAGCATCCTTTGTGAGAAACAGTGAGCATTTTG 51246
QY 241 -----GlnGlyAspSerG1 245
Db 51247 GGGCATAGACACTGACAGTGTATAGAGGCTTTTCCCTCTGCAGGGAGATTCAGG 51306
QY 245 YGlySerLeuMetCysArgAspLysGlyAlaTrpHrLeuAlaGlyValThrSerTr 265
Db 51307 AGGTTACATCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGTGTGACTTCCTG 51366
QY 265 pGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGlnGlySerPr 285
Db 51367 GGGTTGGCTGTGCTGAGGCTGAGAAACAATCTGAGGAAAGTGTATCAGGATCCCC 51426
QY 285 oGlyllePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrG1 305
Db 51427 TGGGATCTTCACAGACATAGTAAAGTGTCTCCCTGGATCCACGAACACATCCAAACTGG 51486
QY 305 yAsn 306
Db 51487 TAAC 51490

```

```

RESULT 8
AC012228/c
LOCUS Homo sapiens chromosome 11 clone RP11-439A13 map 11, *** SEQUENCING
DEFINITION IN PROGRESS ***, 2 ordered pieces.
ACCESSION AC012228
VERSION AC012228.8 GI:22123492
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 180707)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-439A13
JOURNAL Unpublished
AUTHORS

```

```

2 (bases 1 to 180707)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, W., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagsos, B., Hearford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, I., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 180707)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, J., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
QY 100 AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyAspLeuSerGlnThr 119
Db 167186 CATGAAACATGTGCTACTTTGAATTTACTGCTGGAGAGATGACTTAAAGCAGACA 167127
QY 120 AspProGlyGluGlnThrLeuThrIleGluThrValIleHisProHisPheSerThr 139
Db 167126 GACCCAGGAGAGCAACTCTCATTGAACTGTCATCATCATCATCTTCCACC 167067
QY 140 LysLysProMetAspTyAspIleAlaLeuLysMetAlaGlyAlaPheGlnPhe--- 158
|||||

```

```

Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagsos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21362200.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2182
Center clone name: 439_A_13
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 71362: contig of 71362 bp in length
* 71363 71462: gap of 100 bp
* 71463 180707: contig of 109245 bp in length.
Location/Qualifiers
1. 180707
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-439A13"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 47741 a 39327 c 40651 g 52888 t 100 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.67e-66 Length: 180707
Score: 831.00 Matches: 205
Percent Similarity: 30.21% Conservative: 1
Best Local Similarity: 30.06% Mismatches: 1
Query Match: 50.79% Indels: 475
DB: 2 Gaps: 3
US-09-735-713A-2 (1-306) x AC012228 (1-180707)
QY 100 AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyAspLeuSerGlnThr 119
Db 167186 CATGAAACATGTGCTACTTTGAATTTACTGCTGGAGAGATGACTTAAAGCAGACA 167127
QY 120 AspProGlyGluGlnThrLeuThrIleGluThrValIleHisProHisPheSerThr 139
Db 167126 GACCCAGGAGAGCAACTCTCATTGAACTGTCATCATCATCATCTTCCACC 167067
QY 140 LysLysProMetAspTyAspIleAlaLeuLysMetAlaGlyAlaPheGlnPhe--- 158
|||||

```


Db 167066	AAAAACCAATGGACTATGATATGGCCCTTTGAAGATGGCTGGAGCCTTCCAATTTGGT	167007
QY 158	-----	158
Db 167006	AAACATTTGAGGATCATACTGAATGTCTACCCCATGTTCTCTATAACCTCTGTGCCATA	166947
QY 158	-----	158
Db 166946	GCACATTTCTCGAGTATACGTTCTGGGTCTGCACATTTTGAGATAGGAATTTAGGCTGGA	166887
QY 158	-----	158
Db 166886	GAGTCTACCTGAAACTCTTTTGTGTTCTGGATAGCATATACATGGAGTCAGG	166827
QY 158	-----	158
Db 166826	AGAAGGTTTCTTAATTGACAAGTAACCCCTGGAGGTGCCAGGCAGAGCTATCACAACT	166767
QY 158	-----	158
Db 166766	CTCATGAGTCTTTAGGGGACCACCTGGATACCAAGTGTGAGTGGCTTATAGTGCCTC	166707
QY 159	-----GlyHisPheValGlyPro 164	
Db 166706	TCCTCCATGAATGACAGAACCCCTACCTTGCCTTTTCTTTCAGGCCACTTTGTGGGGCCC	166647
QY 165	IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 184	
Db 166646	ATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTTATTGTGTACAACATGCAGGC	166587
QY 185	TrpGlyArgLeuThrGlu-----	190
Db 166586	TGGGGCCGCTTAACGTAAGSTAAGRACTTGTGTGCCATTCTTAAGCAGAGGTCAAGAA	166527
QY 190	-----	190
Db 166526	CTTCTGTGGGCAATCAGAAAAATGGAATTTTAAATCCATTTTAAAAATCGATTGTGCA	166467
QY 190	-----	190
Db 166466	TACATTTCTAAGATAGGAATGTCACATGTACATTAGAAATCAAGCTTGCCAAAGCGA	166407
QY 190	-----	190
Db 166406	CACAAACACAATGATGATCTTCAACTCAGAAAGGTTGAAAGTTCTGCTGGGGCTT	166347
QY 191	-----GlyGlyValLeuSerGlnValLeuGlnGluValAsnLeuPro 204	
Db 166346	CCATCTCTCTGTTTTTCAGGTGGCGCTCTCACAAGTCTTCAGGAAGTGAATCTGCCT	166287
QY 205	IleLeuThrTrpGluGluCysValAlaLeuLeuThrLeuLysArgProIleSerGly 224	
Db 166286	ATTTGACCTGGAGAGGTGTGGCGAGCTCTGTAACTAAGAGGCCATCAGTGGG	166227
QY 225	LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys-----	240
Db 166226	AAGACCTTTCTTGACACAGTTTCTCGATGGAGGAGACGCATGTCAGGTGGGTGGG	166167
QY 240	-----	240
Db 166166	AGCTGGGGGCTGTCAGGCCCCAGGCAAGGATTCCCTTGTTTTGTTTCTTAGTTTTT	166107
QY 240	-----	240
Db 166106	CTTAATTTTCAGTGAATTTAATTTAATTAGAAAGCATATTGGAGGGTAGAATAAGAA	166047
QY 240	-----	240
Db 166046	TAGTCTGAAGAAGCTATAAGGTCAAATAGTTAGTTAAATGTGTGAGAAAAATGTTAGA	165987
QY 240	-----	240
Db 165986	TCCAAACCACTTCCCTAAGAGGCTCACCTGTTATCAGGTGAGGGGCTGCAGGAAGACTGA	165927
QY 240	-----	240
Db 165926	TTCCAGGTAAAGCAGGAATCTGGAAAAGGGTTAGAGTACCAAGGAGGTTTACATCAA	165867
QY 240	-----	240
Db 165866	AATAAAGGCAGCTGGAAAACAATCTGTTTTAGCAGGAAGTGTGATTTAAAGAGGTCTT	165807
QY 240	-----	240
Db 165806	CTGCTTTTCCAAAAGACCATAACAGAGATAATGCCAAGATGAATCGGGGTTCTCC	165747
QY 240	-----	240
Db 165746	CATTCTCTGAAGATGGCTGACTCTTTCCTATTTAGAAATTTTCCAGTTTGTGGATGTGGG	165687
QY 240	-----	240
Db 165686	NACTTTACACTTTTCCCTCCATTCTTCTACTAACCAAGGCCCTGAAGTCAGTCTCATTTGCT	165627
QY 240	-----	240
Db 165626	GGCCAAGACTTCTCTGAGTCTGCAGCATGAGAGGAGAGAAAGTCTGCTGCTTT	165567
QY 240	-----	240
Db 165566	GATGAAAGCCATGGATGAGGAAAAGGAGGTGGGATTTTGGGTTCCAGTGATAGTTGAG	165507
QY 240	-----	240
Db 165506	GATCATATGGGAATTGTCTGGGATCTGTATACACAGGACCCCTGAGTCTAAATTTATAGA	165447
QY 240	-----	240
Db 165446	AGTGGAGATGTCTATTGATCTCTCCAGCATCTTGTGAGGAAACAGTGAGCATTTTG	165387
QY 241	-----GlnGlyAspSerG1 245	
Db 165386	GGCATAGACACTGACAGTGTCTAATAGAGGCTTTTCCCTCTCGAGGAGATTCAGG	165327
QY 245	yGlySerLeuMetCysArgAsnLysGlyAlaIleThrLeuAlaGlyValThrSerTr 265	
Db 165326	AGGTTTCACTCATGTGCGGAATAAGAAAGGGCCCTGACTCTGGCTGGTGTGACTTCCCTG	165267
QY 265	pGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySerPr 285	
Db 165266	GGGTTTGGGCTGTGGTCTGAGGCTGGAGAAACAATGTGAGGAAAGTGCATCAAGGATCCCC	165207
QY 285	oGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrG1 305	
Db 165206	TGGGATCTTCACAGACATTTAGTAAAGTGTCTCCCTGGATCCACGAACATCCAAACTGG	165147
QY 305	yAsn 306	
Db 165146	TAAC 165143	
RESULT 9		
LOCUS	XL081291	
DEFINITION	Xenopus laevis oviductin mRNA, complete cds.	
ACCESSION	U81291	
VERSION	U81291.1 GI:1754713	
KEYWORDS		
SOURCE	Xenopus laevis.	
ORGANISM	Xenopus laevis	
REFERENCE		
AUTHORS	1 (bases 1 to 3689)	
TITLE	Lindsay, L.L., Wieduwilt, M.J., and Hedrick, J.L.	
	Oviductin, the Xenopus laevis oviductal protease that processes egg	
	envelope glycoprotein gp43, increases sperm binding to envelopes,	

AUTHORS Yang, J.C., Lindsey, L.L. and Hedrick, J.L.
 TITLE Direct Submission
 JOURNAL Submitted (06-DEC-1996) Molecular & Cellular Biology, University of California, Davis, CA 95616, USA
 REFERENCE 5 (bases 1 to 4528)

AUTHORS Yang, J.C., Lindsey, L.L. and Hedrick, J.L.
 TITLE Direct Submission

JOURNAL Submitted (24-MAR-1998) Molecular & Cellular Biology, University of California, Davis, CA 95616, USA
 REMARK Nucleotide and amino acid sequence updated by submitter
 COMMENT On or before Oct 11, 2000 this sequence version replaced
 gi:1216404, gi:1754711.

FEATURES

source

Location/Qualifiers

1..4628

/organism="Xenopus laevis"

/db_xref="taxon:8355"

11..4585

/codon_start=1

/product="polyprotein"

/protein_id="AAC24717.1"

/db_xref="GI:2981641"

/translation="MWYSITLLVAAICISAKISNLPPOESPIPETAEKLCGRPE
 IGDEPDLFTSRVGGDAAGVGPWVSLNERHCIGGSIVRKDWVTAARCVIPV
 TEIKVSHMTVIVGSEYDQVMDSQSIPVSHIEPHNPYRGDNGMYDIALVFLSKPII
 FGSVQPTCLPQVGEKIEAGTLCYSSGWRLEENGDLSPVLOEYKLPVVDNNGTCHAVL
 EPIGHVLDPTMLCAGPEGMGACQDGGSPFCRRRSQVWFLAGCVSWGLGCRSW
 CAKQIIRSQSPAFIRSVSLDFLRPKLTGGCSKGRITGKNGTVRYPLSGNSW
 INSVKMLVAOKAKTEIRFLOLDIEDHATCTFDYLSFTVNEKMKIRKCGSTIPSL
 IVRNKTWTFPSDGTGTRGFETQFLAIPKAAACGSAKILKKMKIYSPNPDPY
 PRKTCWIIIEPGENHIVKLAFFNEVGHGCIYDAVEYDGAEEKGLIARLCGYTL
 PLPISSPENTMLIRFKTDMENSYFGFKVFSVPKQFSLPVDPTTISMLHPRAIA
 LDVCGMAPMTKMLVPRIVGEEASPNSPWQVQIFLRTFHCETGAIISQWILTAH
 CIRAEPSYTWYIAGDHNRLNTEQIRNIKTIRIHDNYSNTYDNDIALLYLEPL
 DLNDFRPVCPUEPEEVLTPASVCVWTGNGTADGGOPALGLOQLPILDSIICNTS
 YYSGLTDHMLCAGPPSSKDKACQDGGSLVQNEKEQFSIYGLWSVGGCGRVSK
 PGVTVKELFTWNTQDLOQENALLNSKSVQREKGAAGSGSDVLYVELVGLIA
 SPGPFPYAGSLRWSVIEAAPPISIIKLTLADLSEEDCLNFALILNDSIDKKV
 TRGSLHNKDYIYSSAQGVGHVFGSGDGNRNOGFFLQYWIYVQVDPDLAEFTLSQ
 KPGYSCQDVIYSSGVEIYSPNYLGNYPDLHCQWRITAPGVKIVRVDLELTKTEK
 DVSQCQDLVYNGIGESKOLLGVYGEIRISLSKSGSEITLFTSVSEVSGKFSL
 KYSQDQAGSKOLEENAAVCGCPALDLOIOPESMPGSAELESFNPFSLYSNEMDCOW
 ILYSKGKKLOIEVLYLNLEASPNCSWDYLNIDYGNPNSHLLASLCGHODLVLSN
 GTFVILPHSKSVVDWGFKIQYGAPEPSTLASLDVPSESDSDSESETKQKQD
 CCGNTVDFPMLIYMARSGKIRNLNKGKESGRVVGQQAAPRSPWLVSIOKNSRHHY
 CGGIITINKWLTAAHCEVKNLHRVVVGHDTLVEQNEHAKVYNHSHVHELYMPPSP
 PRNDLLLELPTLLNNSAVICLPDDVTDTMQAECLVAGWGVTDVGGMSLPTKLO
 QAKVPIYVTKKCKDYVSDVTDNNICAGKAGACMGSDGGGPLICKREDRYLYLVGVWS
 WSGCKDQKAPSVYTLTSAFMDWISQHMDT"

mat_peptide

/product="serine protease"

1760..2461

/product="serine protease"

3893..4582

/product="ovo-chymase"

/function="role in the blocking of polyspermy"
 /note="egg chymotrypsin-like protease; located in the
 perivitelline space of the unactivated egg and is released
 at fertilization"

941..1273

/note="encodes a CUB domain"

1295..1633

/note="encodes a CUB domain"

2533..2881

/note="encodes a CUB domain"

2948..3289

/note="encodes a CUB domain"

3341..3700

/note="encodes a CUB domain"

1257 a 1025 c 1166 g 1180 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

4.62e-49

Length:

4628

Score: 626.00 Matches: 128
 Percent Similarity: 57.89% Conservative: 37
 Best Local Similarity: 44.91% Mismatches: 110
 Query Match: 38.26% Indels: 10
 Db: 5 Gaps: 4

US-09-735-713A-2 (1-306) x XLU81290 (1-4628)

Qy 24 GlyIleSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyClnSerLeuValLys 43

Db 104 GTGAAATCTGCAGAGCTCAAGTGTGGACACGCTCGTAAATTTGGTGAAT-----151

Qy 44 ValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyCysSerGlnValGlu 63

Db 152 ---GAACCAGACCTGGAATTCAC-----TCCCGCATTTGGGAGGAGAGATCTGTCT 202

Qy 64 LysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGly 83

Db 203 GTTGGAGGACCAACCATGGACGCTGCGTAAAGCTGAATGAGAGGCATATCTGTGGGGC 262

Qy 84 SerIleValSerProGlnTrpValIleThrAlaAlaHisCysIle-----AlaAsnArg 101

Db 263 AGCATTTGGCGCAAGACATGTTGTAATCGGGCCCACTGTGTCTATCTCTGTGACTGAG 322

Qy 102 AsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspPro 121

Db 323 ATAAAGTGAGCCACATGATCTGATTGTTGGGAATATGACCAGCAAGTTATGGACACT 382

Qy 122 GlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLys 141

Db 383 CAAGACGAGAGATCCCGTATCCCATATAGAGCGCACCAATATACAGGGGATGG 442

Qy 142 ProMetAspTyrAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGlyHisPhe 161

Db 443 ACATGGGCTATGACATGCTCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 502

Qy 162 ValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThr 181

Db 503 GTGCACCCATTTGGCTTCCCGAGGTGGGAGAGAAGATTGAAGCCGGAACCTTGTGTGT 562

Qy 182 ThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnVal 201

Db 563 TCCACGGCTGGGAGGCTTGAAGAAATGGGGACTTGTCCCACTCTCTACAGGAATGA 622

Qy 202 AsnLeuProIleLeuThrTrpGluGluCysValAlaLeuLeuThrLysArgPro 221

Db 623 AAATGCCAGTCGTTGATAATGCGACCTGCCATGCTGTGTGGAGCCCATAGGCATCCC 682

Qy 222 IleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyArgAspAlaCysGln 241

Db 683 GTATTAGATGACACAAATGTTATGCTGATTTCCAGAGAGGAGGATGATGCTGCCAG 742

Qy 242 GlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaThrPheLeuAlaGly 261

Db 743 GCGGATCTCGAGAGACCTTTGTATGACAGCGCGCTGTGGGGTGTGTCTTGGCTGGA 802

Qy 262 ValThrSerTrpGlyCysGlyArgGlyTrp---ArgAsnAsnValArgLysSer 280

Db 803 TGTGTATCTCTGGGCGCTGGCTGTGGAGAACCTGGGGTGGGAGAGATGATGCTGCCAG 862

Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300

Db 863 CAGAGTGGATCCCTGCAATTTTCTCCAGGGTCTCTCAGTGTGTGACTTCTCTGACGCC 922

Qy 301 HisIleGlnThrGly 305

Db 923 CCTAAACTAACTGGA 937

RESULT 12

AX480940

LOCUS AX480940 2627 bp DNA linear PAT 12-AUG-2002

DEFINITION Sequence 36 from Patent WO0246383.

ACCESSION AX480940

Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Unpublished
2 (bases 1 to 169388)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAHL
Center clone name: CH230-388116
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 136807 bases at least Q40
Consensus quality: 141381 bases at least Q30
Consensus quality: 145526 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1102: contig of 1002 bp in length
1103: gap of unknown length
2132: contig of 1030 bp in length
2232: gap of unknown length
3358: contig of 1126 bp in length
3359: gap of unknown length
3458: contig of 1565 bp in length
5023: contig of unknown length
5123: gap of unknown length
5124: gap of unknown length
6663: contig of 1540 bp in length
6763: gap of unknown length
8065: contig of 1302 bp in length
8165: gap of unknown length
9312: contig of 1147 bp in length
9412: gap of unknown length
11980: contig of 2568 bp in length
12080: gap of unknown length
13407: contig of 1327 bp in length
13507: gap of unknown length
13508: contig of 1950 bp in length
15457: contig of unknown length
15557: gap of unknown length
16567: contig of 1010 bp in length
16667: gap of unknown length
17971: contig of 1304 bp in length
18071: gap of unknown length
19740: contig of 1669 bp in length
19741: gap of unknown length
21942: contig of 2102 bp in length
22042: gap of unknown length
23410: contig of 1368 bp in length
23510: gap of unknown length
25146: contig of 1636 bp in length
25246: gap of unknown length
26357: contig of 1111 bp in length
26457: gap of unknown length
26358

26458
28551
28651
29930
30030
31428
31528
33320
33330
33429
34854
34954
36498
36598
39161
39162
39262
40906
41006
44219
44319
46887
46888
49280
49380
49381
51243
51343
53328
53428
55719
55818
59313
59314
59413
61947
62047
62048
64297
64397
66013
66112
68206
68306
68307
71222
71223
71322
71323
74509
74609
77284
77384
77385
80659
80660
80660
80760
84061
84160
86617
86717
86718
89029
89129
92826
92827
92927
97425
97426
97525
100539
100639
100640
106363
106463
106464
111809
111810
111909
115701
115702
115801
115802
121844
121845
121945
130100
130199
130200
137395
137396
143948
143949
144049
151651
28550: contig of 2093 bp in length
28650: gap of unknown length
29929: contig of 1279 bp in length
30029: gap of unknown length
31427: contig of 1398 bp in length
31527: gap of unknown length
33329: contig of 1802 bp in length
33429: gap of unknown length
34854: contig of 1425 bp in length
34954: gap of unknown length
36498: contig of 1544 bp in length
36598: gap of unknown length
39161: contig of 2563 bp in length
39261: gap of unknown length
40905: contig of 1644 bp in length
41005: gap of unknown length
44218: contig of 3213 bp in length
44318: gap of unknown length
46887: contig of 2569 bp in length
46888: gap of unknown length
49280: contig of 2293 bp in length
49380: gap of unknown length
49381: contig of 1862 bp in length
51242: gap of unknown length
51342: gap of unknown length
53327: contig of 1985 bp in length
53427: gap of unknown length
55718: contig of 2291 bp in length
55818: gap of unknown length
59313: contig of 3495 bp in length
59413: gap of unknown length
61947: contig of 2534 bp in length
62047: gap of unknown length
62048: contig of 2249 bp in length
64296: gap of unknown length
64396: gap of unknown length
66012: contig of 1616 bp in length
66112: gap of unknown length
68206: contig of 2094 bp in length
68306: gap of unknown length
68307: contig of 2916 bp in length
71222: gap of unknown length
71322: contig of 3187 bp in length
74509: gap of unknown length
74609: gap of unknown length
77284: contig of 2675 bp in length
77384: gap of unknown length
80659: contig of 3275 bp in length
80759: gap of unknown length
84060: contig of 3301 bp in length
84160: gap of unknown length
86617: contig of 2457 bp in length
86717: gap of unknown length
89029: contig of 2312 bp in length
89129: gap of unknown length
92826: contig of 3697 bp in length
92926: gap of unknown length
97425: contig of 4499 bp in length
97525: gap of unknown length
100539: contig of 3014 bp in length
100639: gap of unknown length
106363: contig of 5724 bp in length
106463: gap of unknown length
111809: contig of 5346 bp in length
111909: gap of unknown length
115701: contig of 3792 bp in length
115801: gap of unknown length
121844: contig of 6043 bp in length
121845: gap of unknown length
121945: contig of 8155 bp in length
130199: gap of unknown length
130200: contig of 7196 bp in length
137395: gap of unknown length
137396: contig of 6453 bp in length
143948: gap of unknown length
144049: contig of 7603 bp in length
151651

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Db 98598 TGATGAGCCTAAAGAAATGGGAAAAGAGCAGCAATTGACAGATTCTGTATGTTTGGGATC 98657

Query Match

Query Match

Query Match:

Length:	942
Matches:	112
Conservative:	46
Mismatches:	98
Indels:	38

[illegible]

Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGlu 300
 ||| |||::: |||:::||||||| :::
 Db 2362 ACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAA 2397

Search completed: March 25, 2003, 03:02:41
Job time : 3507 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 02:04:19 ; Search time 50 Seconds
(without alignments)
815.494 Million cell updates/sec

Title: US-09-735-713A-2
Perfect score: 1636
Sequence: 1 MSLKMLISRNKLLILLGIVF.....IPTDISKVLMSIHEHIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	100.0	306	AAE04733	Human protease hom
2	1632	99.8	306	AAU074760	Human protease PRT
3	1618	98.9	302	AAE04734	Human protease hom
4	1362	83.3	556	AAU82748	Amino acid sequenc
5	950	58.1	182	AAU24275	Toad EST encoded p
6	898	54.9	164	AAE04735	Human protease hom
7	567.5	34.7	913	ABG24246	Novel human diagno
8	567.5	34.7	1576	ABG07870	Novel human diagno
9	567.5	34.7	1576	ABG10218	Novel human diagno
10	567.5	34.7	1576	ABG14588	Novel human diagno

11	567.5	34.7	1576	22	ABG19887	Novel human diagno
12	561.5	34.3	1031	23	ABP60993	Novel human protei
13	489.5	29.9	802	20	AAU41710	Human PRO618 prote
14	489.5	29.9	802	21	ABAB4266	Human PRO618 (UNQ3
15	489.5	29.9	802	21	AAAB2052	Human PRO618 prote
16	489.5	29.9	802	23	AAU82755	Amino acid sequenc
17	486	29.7	787	22	ABB71302	Drosophila melanog
18	485.5	29.7	658	22	AAE06934	Human membrane-ty
19	485.5	29.7	802	22	AAE06933	Human membrane-ty
20	476.5	29.1	235	22	AAE06932	Human membrane-ty
21	467	28.5	452	20	AAU41694	Human PRO382 prote
22	467	28.5	453	22	AAU29055	Human PRO polypept
23	467	28.5	453	22	AAE06935	Human membrane-ty
24	467	28.5	453	23	AAE23020	Human trypsin fami
25	467	28.5	454	21	AAE23246	Tumour associated
26	467	28.5	454	22	AAU68911	Human TAGB-12, pro
27	466	28.5	453	21	AAAB4250	Human PRO382 (UNQ3
28	456.5	27.9	327	21	AAU72093	Human serine prote
29	456.5	27.9	327	23	AAE17921	Human gene 3 encod
30	456.5	27.9	394	23	ABP41994	Human ovarian anti
31	456.5	27.9	454	23	AAE23024	Human trypsin fami
32	456.5	27.9	454	23	AAU82745	Amino acid sequenc
33	454.5	27.8	248	21	AAAB43572	Human cancer assoc
34	454.5	27.8	1128	23	AAU98890	Human protease PRT
35	454.5	27.8	1128	23	AAU82739	Amino acid sequenc
36	446	27.3	243	23	AAU80516	Epithelin-like ser
37	446	27.3	309	23	AAU80531	Epithelin-like ser
38	445	27.2	818	23	AAU82753	Amino acid sequenc
39	444.5	27.2	328	21	AAAB36480	Fusion gene with h
40	444.5	27.2	328	22	AAAB67540	Amino acid sequenc
41	444	27.1	343	23	AAU78547	Human prostaticin pr
42	444	27.1	343	23	ABU07285	Amino acid sequenc
43	444	27.1	414	21	AAE08912	Human secreted pro
44	444	27.1	480	21	AAU80950	Human secreted prote
45	442.5	27.0	417	22	AAE06942	Human hepsin prote

ALIGNMENTS

RESULT 1
AAE04733
ID AAE04733 standard; Protein; 306 AA.
XX
AC AAE04733;
DT 10-SEP-2001 (first entry)
XX
DE Human protease homologue #1.
XX
KW Human; protease homologue; novel human protein; NHP; therapy;
KW pharmacogenomic application; physiological disorder.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 23 /note= "Encoded by CRA"
FT Misc-difference 28 /note= "Encoded by RCT"
XX
PN WO200146407-A1.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US33738.
XX
PR 23-DEC-1999; 99US-0171566.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX

```
DR WPI; 2001-408641743.
DR N-PSDB; AAD09328.
XX
PT Polynucleotide encoding novel human protease homologs, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
PS Claim 2; Page 28-29; 32pp; English.
XX
CC The present sequence is novel human protein (NHP).
CC known as human protease homologue. NHP shares structural similarity
CC with animal proteases, particularly trypsin-like protease such
CC as oviductin, plasminogen activator and human plasma kallikrein
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
CC pharmacogenomic applications. NHP sequences are useful for identifying
CC agonists, antagonists and modulators and also for producing antibodies
CC useful in diagnosis, drug screening, clinical trial monitoring and in
CC treatment of physiological disorders.
XX
SQ Sequence 306 AA;
Query Match 100.0%; Score 1636; DB 22; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.3e-130;
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60
DB 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60
QY 61 QVEKGSYPWQVSLKQKHICGGISVSPQWVITAAHCITANRNIVSTLNTVAGEYDLSQTD 120
DB 61 QVEKGSYPWQVSLKQKHICGGISVSPQWVITAAHCITANRNIVSTLNTVAGEYDLSQTD 120
QY 121 PGEQTLTIETVIIHPHFSTKKPMQYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180
DB 121 PGEQTLTIETVIIHPHFSTKKPMQYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180
QY 181 TTAGWRLTEGVLSQLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240
DB 181 TTAGWRLTEGVLSQLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240
QY 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWNRNVRKSDQSGPIFTDISKVLWIHE 300
DB 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWNRNVRKSDQSGPIFTDISKVLWIHE 300
QY 301 HIQTGN 306
DB 301 HIQTGN 306
RESULT 2
AAU74760
ID AAU74760 standard; Protein; 306 AA.
XX
AC AAU74760;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-20 protein sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.
XX
OS Homo sapiens.
XX
PN WO200198468-A2.
XX
PD 27-DEC-2001.
```

```
XX 13-JUN-2001; 2001WO-US191178.
XX
XX 16-JUN-2000; 2000US-212336P.
XX
PR 22-JUN-2000; 2000US-213955P.
PR
PR 29-JUN-2000; 2000US-215396P.
PR
PR 07-JUL-2000; 2000US-216821P.
PR
PR 14-JUL-2000; 2000US-218946P.
XX
XX (INCYTE GENOMICS INC.
XX
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;
XX Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
XX Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
XX Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
XX Kallick DA;
XX
XX WPI; 2002-090437/12.
XX N-PSDB; ABK12903.
XX
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
XX in the diagnosis, treatment and prevention of gastrointestinal (e.g.
XX gastritis), cardiovascular (e.g. atherosclerosis) and cell
XX proliferative (e.g. cancer) disorders -
XX
XX Claim 1; Page 157-158; 177pp; English.
XX
XX The present invention relates to twenty one new human proteases,
XX referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
XX polypeptides of the invention are useful in the diagnosis, treatment and
XX prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
XX Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
XX myocardial infarction, autoimmune/inflammatory e.g. acquired
XX immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
XX proliferative e.g. cancer, developmental e.g. Duchenne and Becker
XX muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
XX epilepsy and Alzheimer's disease and reproductive e.g. infertility and
XX endometriosis disorders. Numerous other examples of each disorder are
XX given in the specification. The present protein sequence represents
XX the human protease PRTS-20 protein of the invention.
XX
XX Sequence 306 AA;
Query Match 99.8%; Score 1632; DB 23; Length 306;
Best Local Similarity 99.7%; Pred. No. 5.1e-130;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60
DB 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60
QY 61 QVEKGSYPWQVSLKQKHICGGISVSPQWVITAAHCITANRNIVSTLNTVAGEYDLSQTD 120
DB 61 QVEKGSYPWQVSLKQKHICGGISVSPQWVITAAHCITANRNIVSTLNTVAGEYDLSQTD 120
QY 121 PGEQTLTIETVIIHPHFSTKKPMQYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180
DB 121 PGEQTLTIETVIIHPHFSTKKPMQYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180
QY 181 TTAGWRLTEGVLSQLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240
DB 181 TTAGWRLTEGVLSQLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240
QY 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWNRNVRKSDQSGPIFTDISKVLWIHE 300
DB 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWNRNVRKSDQSGPIFTDISKVLWIHE 300
QY 301 HIQTGN 306
DB 301 HIQTGN 306
RESULT 3
```

AAE04734	
ID	AAE04734 standard; Protein; 302 AA.
XX	
AC	AAE04734;
XX	
DT	10-SEP-2001 (first entry)
XX	
DE	Human protease homologue #2.
XX	
KW	Human; protease homologue; novel human protein; NHP; therapy;
KW	pharmacogenomic application; physiological disorder.
XX	
OS	Homo sapiens.
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 19 /note= "Encoded by CRA"
FT	Misc-difference 24 /note= "Encoded by RCT"
FT	FT
XX	WO200146407-A1.
PN	XX
XX	28-JUN-2001.
PD	XX
XX	12-DEC-2000; 2000WO-US33738.
PF	XX
XX	23-DEC-1999; 99US-0171566.
PR	XX
XX	(LEXI-) LEXICON GENETICS INC.
PA	XX
PI	Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
PI	WPI: 2001-408641/43.
DR	N-PSDB; AAD09329.
DR	XX
PT	Polynucleotide encoding novel human protease homologs, useful for
PT	identifying agonist, antagonist or modifiers or for producing
PT	antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT	applications -
XX	Disclosure; Page 29-30; 32pp; English.
XX	XX
CC	The present sequence is novel human protein (NHP),
CC	known as human protease homologue. NHP shares structural similarity
CC	with animal proteases, particularly trypsin-like protease such
CC	as oviductin, plasminogen activator and human plasma kallikrein
CC	precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
CC	pharmacogenomic applications. NHP sequences are useful for identifying
CC	agonists, antagonists and modulators and also for producing antibodies
CC	useful in diagnosis, drug screening, clinical trial monitoring and in
CC	treatment of physiological disorders.
XX	Sequence. 302 AA;
SQ	Query Match 98.9%; Score 1618; DB 22; Length 302;
	Best Local Similarity 100.0%; Pred. No. 7.7e-129;
	Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	5 MLISRNKLIILLGIVFFERKSAALSPLKAPSCGQSLVKVQPNWYFNIFSRILGGSQVEK 64
Db	
	1 MLISRNKLIILLGIVFFERKSAALSPLKAPSCGQSLVKVQPNWYFNIFSRILGGSQVEK 60
QY	65 GSYPMQVSLKORKHICGGGIVSPQWVITAAHCIAANRNIYSTLNVTAGEYDLSQTDPEQ 124
Db	
	61 GSYPMQVSLKORKHICGGGIVSPQWVITAAHCIAANRNIYSTLNVTAGEYDLSQTDPEQ 120
QY	125 TLTETVTHPHFSTKKPMDYDIALKMGAFQFGHEVGPICLPPELREQFEAGFICTAG 184
Db	
	121 TLTETVTHPHFSTKKPMDYDIALKMGAFQFGHEVGPICLPPELREQFEAGFICTAG 180
QY	185 WGRITGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFICTGFPDGGRDACQGS 244
Db	
	181 WGRITGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFICTGFPDGGRDACQGS 240
AAU82748	
ID	AAU82748 standard; Protein; 556 AA.
XX	
AC	AAU82748;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	Amino acid sequence of novel human protease #47.
XX	XX
KW	Human; protease; cancer; immune-related disorder; cardiovascular disease;
KW	neural-associated disease; metabolic disorder; inflammatory disorder;
KW	nervous system disorder; sexual dysfunction; pain; mood disorder;
KW	hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW	viral infection; human immunodeficiency virus; HIV; non-viral infection;
KW	ocular disease; cytostatic; enzyme.
XX	XX
OS	Homo sapiens.
XX	WO200200860-A2.
PN	XX
XX	03-JAN-2002.
PD	XX
XX	26-JUN-2001; 2001WO-US20171.
PF	XX
XX	26-JUN-2000; 2000US-214047P.
PR	XX
XX	(SUG-) SUGEN INC.
PA	XX
PI	Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI	Charyczak G;
XX	WPI: 2002-139913/18.
DR	N-PSDB; ABK31790.
XX	XX
CC	Nucleic acids encoding novel human proteases, useful for useful for
PT	treating diseases and disorders such as cancers, immune-related
PT	diseases and disorders, cardiovascular disease (e.g. restenosis) and
PT	inflammatory disorders -
XX	Claim 6; Fig 2P; 313pp; English.
PS	XX
CC	The present invention relates to the isolation of novel human
CC	proteases, and the nucleic acids encoding them. The sequences of
CC	the invention are useful for treating diseases and disorders such as
CC	cancers (e.g. breast, colon, lung), immune-related diseases and disorders
CC	(e.g. inflammatory diseases and asthma), cardiovascular diseases
CC	(e.g. restenosis and coronary thrombosis), brain or neuronal-associated
CC	diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
CC	disorders (e.g. rheumatoid arthritis and psoriasis), central or
CC	peripheral nervous system diseases, migraines, pain, sexual dysfunction,
CC	mood disorders, attention disorders, cognition disorders, hypotension,
CC	hypertension, psychotic disorders, neurological disorders
CC	(e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
CC	The nucleic acids and polypeptides are also useful for treating viral
CC	infections caused by human immunodeficiency virus (HIV), and non-viral
CC	infections such as ocular disease (e.g. glaucoma) and macular
CC	degeneration. AAU82702-AAU82760 represent the novel human proteases of
CC	the invention.
XX	Sequence 556 AA;
SQ	Query Match 83.3%; Score 1362; DB 23; Length 556;

Best Local Similarity 85.7%; Pred. No. 6.7e-107;
Matches 263; Conservative 8; Mismatches 28; Indels 8; Gaps 2;
Qy 1 MSLKMLISRNKLIILLGIVFFERKSAALSLPKAPSCGQSLVKVQPNYFNIFSRILGGS 60
Dy 1 MSLKMLISRNKLIILLGIVFFERKSAALSLPKAPSCGQSLVKVQPNYFNIFSRILGGS 60
Qy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120
Dy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120
Qy 121 PGEQTLTETVLIHPHFSTKPKMDYDIALKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180
Dy 121 PGEQTLTETVLIHPHFSTKPKMDYDIALKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180
Qy 181 TTAGHGRTEGVLSQVLEQVNLPLTWEBCVAALLTLKRPISGKTFCTGFPDGGDRAC 240
Dy 181 TTAGHGRTEGVLSQVLEQVNLPLTWEBCVAALLTLKRPISGKTFCTGFPDGGDRAC 240
Qy 241 QDGSGLMCRNKKGAWTLAGVTSMLGCGGRWNRNVRKSDQSP-----GIFTDISK 293
Dy 241 QDGSGLMCRNKKGAWD-SGWSIWEAQVGGSLSSRSRPSLGNKVRCLTNFFKKLAG 299
Qy 294 VLSWIHE 300
Dy 300 CGTWCSE 306

RESULT 5
AAM24275
ID AAM24275 standard; Protein: 182 AA.
XX AC AAM24275;
XX DT 12-OCT-2001 (first entry)
XX DE Toad EST encoded protein SEQ ID NO: 1800.
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX KW diagnostics; forensic test; gene mapping; genetic disorder;
XX KW biodiversity; gene therapy; nutrition.
XX OS Xenopus laevis.
XX PN WO200154477-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US02687.
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 17-JUL-2000; 2000US-0617746.
XX PR 03-AUG-2000; 2000US-0631451.
XX PR 15-SEP-2000; 2000US-0663870.
XX PA (HYSE-) HYSEQ INC.
XX Qy Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX DR WPI; 2001-476164/51.
XX DR N-PSDB; AAH98934.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX PT antibodies and research use -
XX PS Claim 20; Page 1178-1179; 1275pp; English.
XX CC The present invention provides the protein and coding sequences of novel
XX CC proteins from a variety of organisms, including human, dog, cat, horse,
XX CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX CC urchin and tomato. These were derived from expressed sequence tags (ESTs)

CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.
XX
SQ Sequence 182 AA;
Query Match 58.1%; Score 950; DB 22; Length 182;
Best Local Similarity 99.5%; Pred. No. 1.3e-72;
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MSLKMLISRNKLIILLGIVFFERKSAALSLPKAPSCGQSLVKVQPNYFNIFSRILGGS 60
Dy 1 MSLKMLISRNKLIILLGIVFFERKSAALSLPKAPSCGQSLVKVQPNYFNIFSRILGGS 60
Qy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120
Dy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120
Qy 121 PGEQTLTETVLIHPHFSTKPKMDYDIALKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180
Dy 121 PGEQTLTETVLIHPHFSTKPKMDYDIALKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180
Qy 181 TT 182
Dy 181 TT 182
RESULT 6
AAE04735
ID AAE04735 standard; Protein: 164 AA.
XX AC AAE04735;
XX DT 10-SEP-2001 (first entry)
XX DE Human protease, homologue #3.
XX KW Human; protease homologue; novel human protein; NHP; therapy;
XX KW pharmacogenomic application; physiological disorder.
XX OS Homo sapiens.
XX PN WO200146407-A1.
XX PD 28-JUN-2001.
XX PF 12-DEC-2000; 2000WO-US33738.
XX PR 23-DEC-1999; 99US-0171566.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX DR WPI; 2001-408641/43.
XX DR N-PSDB; AAD09330.
XX PT Polynucleotide encoding novel human protease homologs, useful for
XX PT identifying agonist, antagonist or modifiers or for producing
XX PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
XX PT applications -
XX PS Disclosure; Page 30-31; 32pp; English.
XX CC The present sequence is novel human protein (NHP),
XX CC known as human protease homologue. NHP shares structural similarity
XX CC with animal proteases, particularly trypsin-like protease such
XX CC as oviductin, plasminogen activator and human plasma kallikrein
XX CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
XX CC pharmacogenomic applications. NHP sequences are useful for identifying
XX CC agonists, antagonists and modulators and also for producing antibodies
XX CC useful in diagnosis, drug screening, clinical trial monitoring and in

CC treatment of physiological disorders.

XX Sequence 164 AA;

Query Match 54.9%; Score 898; DB 22; Length 164;

Best Local Similarity 100.0%; Pred. No. 2.9e-68;

Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 MDYDIALKMGAFQFGHFGVPTICLPPELREQFEAGFICTTAGWRLTEGGVLSQVLEVN 202

DB 1 MDYDIALKMGAFQFGHFGVPTICLPPELREQFEAGFICTTAGWRLTEGGVLSQVLEVN 60

QY 203 LPILTWEECVAALLTLKRPIISGKTFELCTGPPDGGDRACQDGGSLMCRNKKGAWLAV 262

DB 61 LPILTWEECVAALLTLKRPIISGKTFELCTGPPDGGDRACQDGGSLMCRNKKGAWLAV 120

QY 263 TSWGLCGGRWRNVRKSDQSGIETDTSKVLWHEHIQTGN 306

DB 121 TSWGLCGGRWRNVRKSDQSGIETDTSKVLWHEHIQTGN 164

RESULT 7

ABG24246

ID ABG24246 standard; Protein: 913 AA.

XX ABG24246;

XX

AC ABG24246;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #24237.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

XX WO200175067-A2.

PN

PD 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US08631.

PF

XX 31-MAR-2000; 2000US-0540217.

XX

PR 23-AUG-2000; 2000US-0649167.

XX

XX (HYSE-) HYSEQ INC.

PA

XX Drmanac RT, Liu C, Tang YT;

PI

DR WPI: 2001-639362/73.

XX

DR N-PSDB; AAS88433.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

XX Claim 20; SEQ ID No 54605; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 913 AA;

Query Match 34.7%; Score 567.5; DB 22; Length 913;

Best Local Similarity 41.2%; Pred. No. 1.8e-39;

Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;

QY 37 CGOSLVKV---QPNYFNIESRILGGSQVEKGYPMQVSLKQKQKHTCGGSIYSPQWVIT 93

DB 9 CGIRMVNMKSKEPAVGSRRFFSRISRNSTVTGHPMVQVSLKSDHFFCGSLIQEDRVVT 68

QY 94 AAHCII--ANRNIVSTLVNTAGEYDLSQTPGEQTLTETVIIHPHFTKKRPMYDIALLK 151

DB 69 AAHCLDSLSEKQLKNITVTSGEYSLFKQKQEQNIPYSKIITHEPYSNREYMSFDIALLY 128

QY 152 MAGAFQGFHVGPICLPELREQFEAGFICTTAGWRLTEGGVLSQVLEVNLPILTWEBC 211

DB 129 LKHKVFQGNVQPICLPDSDDKVEPGILCLSSGWGKISKTSYXSNVLQEMELPIMDDRAC 188

QY 212 VAALLTLKRPIISGKTFELCTGPPDGGDRACQDGGSLMCRNKKGAWLAVTWSGLGCCR 271

DB 189 NTVLKMNLPLGRTMLCAGFPDWMGMDACQDGGPLVCRGGGIIWLAGITTSWAGCAG 248

QY 272 G---WRNVKSKDQSGPIETDISKVLWHEHIQTG 305

DB 249 GSVPRNVNHVKA---SLGIFSKVSELMDFITQNLFTG 282

XX

XX RESULT 8

XX ABG07870

XX ID ABG07870 standard; Protein: 1576 AA.

XX

XX AC ABG07870;

XX

XX 13-FEB-2002 (first entry)

XX

XX Novel human diagnostic protein #7861.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

XX Homo sapiens.

XX

XX WO200175067-A2.

XX

XX 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US08631.

XX

XX 31-MAR-2000; 2000US-0540217.

XX

XX 23-AUG-2000; 2000US-0649167.

XX

XX (HYSE-) HYSEQ INC.

XX

XX Drmanac RT, Liu C, Tang YT;

XX

XX WPI: 2001-639362/73.

XX

XX N-PSDB; AAS72057.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

XX Claim 20; SEQ ID No 38229; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1576 AA;

Query Match 34.7%; Score 567.5; DB 22; Length 1576;
Best Local Similarity 41.2%; Pred. No. 3.3e-39;
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;
QY 37 CQOSLVK---QPNWYNFISRLGSGQVEKSGYPQVSLKQKQKHICGGSTVSPQWVIT 93
|| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 848 CGIRMVNKKSEKPEAVGSRFFSRISWRNSTVTGHPQVSLKSDHFCGSLIQEDRVVT 907
|||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
QY 94 AAHCI--ANRNIVSTLNVTAGYDLSQTPGQTLTIETVIHPHFSTKPKMDYDIALLK 151
|||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 908 AAHCLDSLSEKQLKNITVTSGEYSLFQDKQEQNPVSKIIITPEYNSREYMSPDIALY 967
QY 152 MAGAFQGHFVGPICLPDLREQFAGFICTTAGWRLTEGGVLSQVLQEVNLPILTWECC 211
||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 968 LKHVKFGNAVQPICLPDSDDKVEPILCLSSGKGKISKTSYSNVLQEMELPIMDDRAC 1027
|| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
QY 272 G---WRNVKRSQDQSGPGIFTDISKVLWIHEHIOTG 305
| ||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 1088 GSVPRVNNHVKA---SLGIFSKVSELMDFITQNLFTG 1121

RESULT 9

ABG10218
ID ABG10218 standard; Protein: 1576 AA.

XX AC ABG10218;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #10209.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS74405.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity

PS Claim 20; SEQ ID No 40577; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1576 AA;

Query Match 34.7%; Score 567.5; DB 22; Length 1576;
Best Local Similarity 41.2%; Pred. No. 3.3e-39;
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;

QY 37 CQOSLVK---QPNWYNFISRLGSGQVEKSGYPQVSLKQKQKHICGGSTVSPQWVIT 93
|| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 848 CGIRMVNKKSEKPEAVGSRFFSRISWRNSTVTGHPQVSLKSDHFCGSLIQEDRVVT 907
|||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
QY 94 AAHCI--ANRNIVSTLNVTAGYDLSQTPGQTLTIETVIHPHFSTKPKMDYDIALLK 151
|||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 908 AAHCLDSLSEKQLKNITVTSGEYSLFQDKQEQNPVSKIIITPEYNSREYMSPDIALY 967
QY 152 MAGAFQGHFVGPICLPDLREQFAGFICTTAGWRLTEGGVLSQVLQEVNLPILTWECC 211
||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 968 LKHVKFGNAVQPICLPDSDDKVEPILCLSSGKGKISKTSYSNVLQEMELPIMDDRAC 1027
|| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
QY 212 VAALLTLKRPISGKTFLCTGPDGGRDACQDGSGLMCRNKKGAWTLAGVTSWLGCGR 271
| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 1028 NTVLKSMLPPLGRMTLCAGFPDWMGMDACQDGSGLVCRGGGIWILAGITSWVAGCAG 1087
|| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
QY 272 G---WRNVKRSQDQSGPGIFTDISKVLWIHEHIOTG 305
| ||| : : : : : |||| : : : : : |||| : : : : : |||| : : : : :
DB 1088 GSVPRVNNHVKA---SLGIFSKVSELMDFITQNLFTG 1121

RESULT 10

ABG14588
ID ABG14588 standard; Protein: 1576 AA.

XX AC ABG14588;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #14579.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI
XX
DR N-PSDB; AAS78775.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 20; SEQ ID No 44947; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1576 AA;
Query Match 34.7%; Score 567.5; DB 22; Length 1576;
Best Local Similarity 41.2%; Pred. No. 3.3e-39;
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;
QY 37 CGOSLVK---OPWNYFNIFSRILGSGQVEKGYPMQVSLKQKQKHICGSGIVSPQWVIT 93
DB 848 CGIRMYNMKKEPVGCSRRFSRISWRNSTVTGHPMQVSLKSDHHFCCGSLIQEDRVVT 907
QY 94 AARCI--ANRNVSTLNTVAGEYDLSQTDPGQTLTIETVTHPHFSTKPKMDYDIALK 151
DB 908 AAHCLDSLSEKQLKNITVTSGEYSLFQDKQEQNIPVSKIIITHPEVNSREYKSPDIALLY 967
QY 152 MAGAFQFHGVPCICLPRLRQEFAGICTTAGWGRLEGGVLSQVLOEVNLPILTWEBC 211
DB 968 LKHVKFQGNVAPQICLPDSDRRVPEGILCLSSGCKGKISTSEYSNVLQEMELPIMDDRAC 1027
QY 212 VAALLTLKRPISGKFLCTGFFDGGDRDAGCGSGSLMCRNKKGAWTLAGVTSWGLGCCR 271
DB 1028 NTVLMSNLPPIGRITMLCAGFPDWGMDACQSGDGGPLVCRGGGIWILAGITTSWAGCAG 1087
QY 272 G---WRNNVRKSDQSGPGIFTDISKVLWSIHEHIQTG 305

DB 1088 GSPVRRNHVKA---SLGIFSKVSELMDFITQNLFTG 1121
RESULT 11
ABG19887
ID ABG19887 standard; Protein; 1576 AA.
XX
XX ABG19887;
XX
XX 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #19878.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX
XX
PI
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS84074.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 50246; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1576 AA;
Query Match 34.7%; Score 567.5; DB 22; Length 1576;
Best Local Similarity 41.2%; Pred. No. 3.3e-39;
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;
QY 37 CGOSLVK---OPWNYFNIFSRILGSGQVEKGYPMQVSLKQKQKHICGSGIVSPQWVIT 93
DB 848 CGIRMYNMKKEPVGCSRRFSRISWRNSTVTGHPMQVSLKSDHHFCCGSLIQEDRVVT 907

XX PD 16-SEP-1999.
XX PF 08-MAR-1999; 99MO-US05028.
XX PF 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-MAY-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.

PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.

(GETH) GENENTECH INC.

Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;

WPI: 1999-551358/46.

N-PSDB; AA234033.

New secreted and transmembrane polypeptides and their polynucleotides,
useful for treating blood coagulation disorders, cancers and cellular
adhesion disorders -

Claim 12; Fig 63; 530pp; English.

The present invention describes secreted and transmembrane polypeptides
and their polynucleotides. The nucleotide sequences are useful as
sources of probes, primers, for chromosome mapping, and for generation
of antisense sequences. They can also be used to create transgenic
animals. The proteins can be used to treat a variety of diseases and
disorders, depending on their function. Diseases that may be treated
include blood coagulation disorders, cancers and cellular adhesion
disorders. They may also be used to raise antibodies. AA233891 to
AA234338, and AA241685 to AA241774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
invention.

XX SQ Sequence 802 AA;

Query Match 29.9%; Score 489.5; DB 20; Length 802;

Best Local Similarity 39.3%; Pred. No. 6e-33;

Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps 5;

QY 54 SRLGGQVEKGYSPWQVSLKQKOKHICGGSGIVSPQWVITAAHCIAHNRNIVSTL--NVTA 111

Db 566 SRIVGGAVSSEGEWPMQASLQVRGRHICGGGALIAHNRVITAAHCFOEDSMASVTLVTVFL 625

QY 112 GEYDLSTQDPGEQTLTETVLIHPHFSTKKPMYDIALKMGAFQFGHFGVPTCLPELR 171

Db 626 GKVMQNSRWPFGEVSFKVSRLLLHP-YHEEDSHDIDVALLQLDHPVVRSAARVPCLPARS 684

QY 172 EQEAGFICTAGWRLTEGGVLSQVLQEVNLPILITWEECVAAALLTLKRPTISGKTLCTG 231

Db 685 HFTEPGLHCWITGVALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTP---RLCAG 740

QY 232 FPDGGRDACDGGSGSLMCRNKKCAWTLAGVTSWGLCGGRGNRNVRKSDGSP---GIF 288

Db 741 YRKCKDACDGGSGGLVCKALSGRWFLAGLVSWGLGCGR-----PNYFGVY 787

QY 289 TDISKVLWSIHE 300

Db 788 TRITGVISWIOQ 799

RESULT 14	
AAB44266	
ID	AAB44266 standard; Protein; 802 AA.
XX	
AC	AAB44266;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human PRO618 (UNQ354) protein sequence SEQ ID NO:169.
XX	
KW	Human; secreted protein; transmembrane protein; PRO; EST; cytosstatic;
KW	expressed sequence tag; detection; cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO200053756-A2.
XX	
PD	14-SEP-2000.
XX	
PF	18-FEB-2000; 2000WO-US04341.
XX	
PR	08-MAR-1999; 99WO-US05028.
PR	12-MAR-1999; 99US-0123957.
PR	29-MAR-1999; 99US-0126773.
PR	21-APR-1999; 99US-0130232.
PR	28-APR-1999; 99US-0131445.
PR	14-MAY-1999; 99US-0134287.
PR	23-JUN-1999; 99US-0141037.
PR	26-JUL-1999; 99US-0145698.
PR	29-OCT-1999; 99US-0162506.
PR	30-NOV-1999; 99WO-US28313.
PR	02-DEC-1999; 99WO-US28551.
PR	02-DEC-1999; 99WO-US28565.
PR	16-DEC-1999; 99WO-US30095.
PR	30-DEC-1999; 99WO-US31243.
PR	30-DEC-1999; 99WO-US31274.
PR	05-JAN-2000; 2000WO-US00219.
PR	06-JAN-2000; 2000WO-US00277.
PR	06-JAN-2000; 2000WO-US00376.
XX	
XX	(GETH) GENENTECH INC.
PA	
XX	
PI	Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI	Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI	Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI	Klajavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI	Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX	
DR	WPI; 2000-611443/58.
DR	N-PSDB; AAC78494.
XX	
PT	Novel PRO polypeptides and polynucleotides used in detection methods,
PT	to target bioactive molecules to specific cells, and to modulate
PT	cellular activities -
XX	
XX	Claim 12; Fig 63; 636pp; English.
XX	
CC	AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC	sequence tag) sequences which encode secreted or transmembrane PRO
CC	polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC	activity. The polynucleotides and polypeptides can be used for detecting
CC	the presence of PRO polypeptides in samples, for linking bioactive
CC	molecules to cells and for modulating biological activities of cells,
CC	using the polypeptides for specific targeting. The polypeptide targeting
CC	can be used to kill the target cells, e.g. for the treatment of cancers.
CC	The polypeptide pairs provide specific targeting of bioactive molecules
CC	to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC	the isolation of the PRO polynucleotide sequences.
XX	
SQ	Sequence 802 AA;
Query Match	29.9%; Score 489.5; DB 21; Length 802;
Best Local Similarity	39.3%; Pred No. 6e-33;

Search completed: March 25, 2003, 03:57:21
Job time : 52 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	688.5	42.1	1004	2	T30338	oviductin (EC 3.4. polyprotein - Afri
2	626	38.3	1524	2	T30337	serine proteinase
3	486	29.7	786	1	A47547	plasma kallikrein
4	466	28.5	638	1	KQHUP	plasma kallikrein
5	464	28.4	638	1	KQMSPL	plasma kallikrein
6	447.5	27.4	638	1	KQRTPL	proctasin (EC 3.4. 1
7	444	27.1	343	1	A57014	trypsin (EC 3.4. 21
8	442.5	27.0	417	1	S00845	hepsin (EC 3.4. 21
9	442.5	27.0	1035	1	A43090	enteropeptidase (E
10	438.5	26.8	416	1	S33777	hepsin (EC 3.4. 21
11	438.5	26.8	1034	1	A53663	enteropeptidase (E
12	433	26.5	593	2	S45281	coagulation factor
13	431.5	26.4	1019	1	A56318	enteropeptidase (E
14	421	25.7	855	1	JC7731	membrane-bound arg
15	421	25.7	855	2	JC7775	membrane type-seri
16	420.5	25.7	1047	2	A56617	masquerade precurs
17	419.5	25.6	237	1	TRCY1	trypsin (EC 3.4. 21
18	419	25.6	264	2	I38136	chymotrypsin-like
19	418	25.6	625	1	KFHU1	coagulation factor
20	415.5	25.4	275	2	S40007	trypsin (EC 3.4. 21
21	414.5	25.3	615	1	KFHU12	coagulation factor
22	413.5	25.3	263	2	A21195	chymotrypsin (EC 3
23	409	25.0	263	2	A31299	chymotrypsin (EC 3
24	409	25.0	275	2	S40005	trypsin (EC 3.4. 21
25	408.5	25.0	415	1	A34170	acrosin (EC 3.4. 21
26	408.5	25.0	603	2	S28941	coagulation factor
27	407.5	24.9	274	2	S35339	trypsin (EC 3.4. 21
28	406	24.8	254	1	TRWV3Y	trypsin-like prote
29	405	24.8	265	2	T15451	hypothetical prote

[illegible]

RESULT 7
A57014
prostatin (EC 3.4.21.-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 18-Jun-1999
C:Accession: A57014; A54866
R:Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 270, 13483-13489, 1995
A:Title: Molecular cloning, tissue-specific expression, and cellular localization of hum
A:Reference number: A57014; MUID:95286644; PMID:7768952
A:Accession: A57014
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-343 <RES>
A:Cross-references: GB:I41351; NID:9862304; PIDN:AAC41759.1; PID:9862305
A:Experimental source: prostate
A:Note: parts of this sequence were determined by protein sequencing
R:Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 269, 18843-18848, 1994
A:Title: Prostatin is a novel human serine proteinase from seminal fluid. Purification,
A:Reference number: A54866; MUID:94308140; PMID:8034638
A:Accession: A54866
A:Molecule type: protein
A:Residues: 45-64 <YUA>
C:Genetics:

[illegible]

10

A:Accession: A43090
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBU
A:Molecule type: mRNA
A:Residues: 1-1035 <KIT>
A:Cross-references: GB:U09859; NID:g746410; PIDN:AAB40026.1; PID:g746411
A:Experimental source: small intestine
A:LaValle, E.R.; Rehmentulla, A.; Racie, L.A.; DiBlasio, E.A.; Ferenz, C.; Grant, K.L.; J. Biol. Chem. 268, 23311-23317, 1993
A:Title: Cloning and functional expression of a cDNA encoding the catalytic subunit of h
A:Reference number: A48874; MUID:94043122; PMID:8226855
A:Accession: A48874
A:Molecule type: mRNA
A:Residues: 801-1035 <LAV>
A:Cross-references: GB:L19663; NID:g416131; PIDN:AAAL6035.1; PID:g416132
A:Note: parts of this sequence, including the amino end of the mature protein, were conf
R:Light, A.; Janska, H.
J. Protein Chem. 10, 475-480, 1991
A:Title: The amino-terminal sequence of the catalytic subunit of bovine enterokinase.
A:Reference number: A61436; MUID:92189715; PMID:1799406
A:Accession: A61436
A:Molecule type: protein
A:Residues: 801-807 'Y', 809-827 <LIG>
C:Comment: The mechanism of association with the membrane of the intestinal brush border
embrane attachment using a signal-anchor sequence.
C:Comment: Conversion from membrane-bound to soluble forms may involve further processin
C:Complex: mature enteropeptidase is variously reported to contain two (heavy and light)
Lfide linked
C:Function:
A:Description: cleaves propeptide from trypsinogen to produce active trypsin
A:Pathway: intestinal digestive hydrolase cascade
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding rep
C:Keywords: glycoprotein; hydrolase; intestine; serine proteinase; transmembrane protein
F:23-38/Domain: transmembrane #status predicted <TM>
F:52-117/Product: enteropeptidase mini chain #status predicted <MC>
F:118-800/Product: enteropeptidase heavy chain #status predicted <HC>
F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:358-520/Domain: MAM homology <MAM>
F:542-647/Domain: C1r/C1s repeat homology <C1r>
F:659-693/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:694-799/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
F:801-1030/Product: enteropeptidase light chain #status predicted <LC>
F:801-1030/Product: trypsin homology <TRY>
F:116,147,170,194,233,263,264,404,456,486,519,550,646,698,722,741,762,864,903,965/Bindin
F:788-912,826-842,926-993,957-972,983-1011/Disulfide bonds: #status predicted
F:841,892,987/Active site: His, Asp, Ser #status predicted

Query Match 27.0%; Score 442.5; DB 1; Length 1035;
Best Local Similarity 34.9%; Pred. No. 7.4e-32;
Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;

QY 36 SCGSLV--KVOPNNYIFSRILGSGVEKSYPMQVSLKQKQKHCIGSGIVSPQWVIT 93
DB 787 SCGKLLVTQVEVP-----KIVGSDSRREGAPWVWVYFDQVCGASLSVSRDWLVS 838
QY 94 AAHCIAANRN1-VSTLNVTAGEYDLSQ-TDPGEOTLTITETVLIHPFSTKPKMDYDIALLK 151
DB 839 AAHCYGRNNEPSKWKAVGLHNASLTSQIETRLDQIVINPHYN-KRKNNDIAMMH 897
QY 152 MAGAFQGHFVGPICLPDLREQPEAGICTAGNGRTEGGVLSQVLEQVNLPLTWTBEC 211
DB 898 LEMKNVNTDYIQTICLPDENQVPPGRICSTAGWALTYQGSTADVLQEQADVPLLSNEK 957
QY 212 VAAL---LTLKRPISKTKTLCTGFPDGRDACOGDSGSGSLMCKNKKGAWTLACVTSWGL 267
DB 958 QOOMPEYNIT-----ENMVCAEYAGVDSCOGDSGGLPMCO-ENNRLLACVTSFGY 1009
QY 268 GCGRGNRNVRKSDQSGPGIFTDTSKVLSTHEHI 302
DB 1010 QCALPNR-----PGYARVPRETIWISFL 1034

RESULT 10
S33777

hepsin (EC 3.4.21.-) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: S33777; S32013
R:Farley, D.; Raymond, F.; Nick, H.
Biochim. Biophys. Acta 1173, 350-352, 1993
A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteina-
A:Reference number: S33777; MUID:93305733; PMID:8318546
A:Accession: S33777
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <FAR>
A:Cross-references: EMBL:X70900; NID:g57928; PIDN:CAA50256.1; PID:g57929
C:Superfamily: hepsin; trypsin homology
C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F:22-44/Domain: transmembrane #status predicted <TM>
F:162-399/Domain: trypsin homology <TRY>
F:187-203,290-359,321-337,348-380/Disulfide bonds: #status predicted
F:202,256,352/Active site: His, Asp, Ser #status predicted

Query Match 26.8%; Score 438.5; DB 1; Length 416;
Best Local Similarity 33.0%; Pred. No. 5.4e-32;
Matches 102; Conservative 48; Mismatches 112; Indels 47; Gaps 11;

QY 11 KLILLGIVFFERKSAALSPLKAPSCGSLVKVQPNWYFNIFSRILGSGQVEKSGYPMQ 70
DB 129 RLDDVISVCDPCGRFLTAT---QDCGRKLPV-----DRIVGGQDSSLGRWPWQ 176
QY 71 VLSKQRKHICGSGIVSPQWVITAAHCIAANRN-IVSTLNVTAGEYDLSQTDPGEQTLTIE 129
DB 177 VSLRYDCHLTCGSLGSDWVLTAAHCFPERNRLSRVLRVFAQA--VARTSPHAVQLGVQ 234
QY 130 TVIHPHF-----STKPKMDYDIALKMGAFQGHFVGPICLPDLREQPEAGICTAG 184
DB 235 AVIYHGGLPFRDPTDITENSNDIALVHLSSSLPTETIQVPCLPAAGQALVDGKVCVTG 294
QY 185 WGLRTEGVLSQLQEVNLPILTWEQVAAALLTKRP-----TSGKTFELCTGFPDGRD 238
DB 295 WGNTOFGQAVVLOEARVPIISNEVC-----NSPDFYGNQIKPKMF-CAGYPEGID 346
QY 239 ACQGDGSGSLMCKRNK---KGAWTLAGVTSWGLGCGRGNRNVRKSDQSGPGIFTDISKVL 295
DB 347 ACQGDGSGHFCEDRIGTSRWRLCGIVSWGTGAL-----ARK-----PGVTKVIDER 396
QY 296 SWIHEHTQT 304
DB 397 EWIFQAIKT 405

RESULT 11
A53663
enteropeptidase (EC 3.4.21.9) precursor - pig
N:Alternate names: enterokinase
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Oct-1994 #sequence_revision 09-Aug-1996 #text_change 18-Jun-1999
C:Accession: A53663
R:Watsushima, M.; Ichinose, M.; Yahagi, N.; Kakei, N.; Tsukada, S.; Kuroka
J. Biol. Chem. 269, 19976-19982, 1994
A:Title: Structural characterization of porcine enteropeptidase.
A:Reference number: A53663; MUID:94327548; PMID:8051081
A:Accession: A53663
A:Molecule type: mRNA
A:Residues: 1-1034 <MAT>
A:Cross-references: GB:D30799; NID:g505122; PIDN:BAA06459.1; PID:g505123
A:Note: parts of this sequence, including the amino ends of three chains isolated fr
C:Comment: The mechanism of association with the membrane of the intestinal brush bo
ated below) or with amino-terminal myristoylation of the heavy chain.
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and li
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms in
C:Function:
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
A:Pathway: intestinal digestive hydrolase cascade
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding

C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
F:22-38/Domain: transmembrane #status predicted <TM>
F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>
F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:357-519/Domain: MAM homology <MAM>
F:541-646/Domain: C1r/C1s repeat homology <C1R>
F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:693-798/Domain: scavenger receptor cysteine-rich domain homology <SRC>
F:800-1034/Product: enteropeptidase light chain #status predicted <LCH>
F:800-1029/Domain: trypsin homology <TRY>
F:116,147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,902,96
F:787-911,825-841,925-992,956-971,982-1010/Disulfide bonds: #status predicted
F:840,891,986/Active site: His, Asp, Ser #status predicted

Query Match 26.8%; Score 438.5; DB 1; Length 1034;
Best Local Similarity 30.3%; Pred. No. 1.7e-31;
Matches 101; Conservative 62; Mismatches 93; Indels 77; Gaps 11;

QY 15 LLGI-----VFFERKSAALSPLKAP-----SCQSILV 42
Db 733 LLGLGTGSSMFFSSGGGPFVKLTAPNGSLILTAEQCFEDSLILQCNHKSCGKKQV 792

QY 43 --KVOPWNYFNFSRLGSGQVEKSGYKQVSLKQKHICGGSIVSQWVITAHCIAN 100
Db 793 AOEVSF-----KIVGNDREGAWPMVALYNYQLCCGASLVSRDLVSAARCVYG 844

QY 101 RNI-----VSTLVNVTAGEYDLSQTDGQETLTITETIIHPHFSTKPKMDYDIALKKMA 153
Db 845 RNLEPSKWKAILGLHMTS-----NLTSPOIVTRLIDEIVINPHYNNRR-KDSDIAMHLE 898

QY 154 GAFQGHFVGPICLDELREQFAGICTAGWGRTEGGVLSQVLQEVNLPILTWEBCVA 213
Db 899 FKNVYTDYIQCLPEENQVFPFGICISAGKVIYQGSFADILQEQADVPILLSNEKQQ 958

QY 214 AL----LTLKRPISGKTLCTGFPDGGDACQDGGSLMCRKKGAWTLAGVTSWGLGC 269
Db 959 QMPEYNIT-----ENMWCAGYEGGIDSCQDGGSLMCL-ENNRWLLAGVTSFGYQC 1010

QY 270 GRGWRNVKRSQDGSFGIFTDISKVLSWIHEH 302
Db 1011 ALPNR-----PGVYARVPKFTWISQL 1033

RESULT 12
S45281
coagulation factor XIIfa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): comp
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SH1>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70 a
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as Gl
Is, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:861210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16,'X',18-19;525-550 <FUJ>
C:Superfamily: coagulation factor XII; EGF
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma; s
F:37-78/Domain: fibronectin type II repeat homology <IF2>
F:88-120/Domain: EGF homology <EGF>

F:125-160/Domain: fibronectin type I repeat homology <FB1>
F:207-287/Domain: kringle homology <KRG>
F:350-587/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted

Query Match 26.5%; Score 433; DB 2; Length 593;
Best Local Similarity 36.1%; Pred. No. 2.7e-31;
Matches 101; Conservative 45; Mismatches 94; Indels 40; Gaps 12;

QY 37 CGOSLVKVPWNYFNFSRLGSGQVEKSGYKQVSLKQKHICGGSIVSQWVITAHAH 96
Db 336 CGQRLK-----WLSLNKRVGLVALGPAHYIAALYWDQ-HFCAGSLIAPCWVLTAAH 389

QY 97 CIANRNVSTLVNVTAGEYDLSQTDGQETLTITETIIHPHFSTKPKMDY--DIALKKMA 154
Db 390 CLQNPAPKELTVLGDQRHNSCEQCQLAVRDYRLHEAFS--PITYQHDALVLRLOE 446

QY 155 AFQ--FGH---FVGPICLPE--LREQFAGFICTTAGWGRLEGGVLSQVLQEVNLPILT 207
Db 447 SADGCAHPSPFPVQCLPSTAARPAESAACEVAGMVGHQFEGGEYSFLQEAQVPLID 506

QY 208 WEECVAAALLTLKRPISGKTF----LCTGFPDGGDACQDGGSLMCRNK--KGAWTLA 261
Db 507 PORCSAP-----DVHGAAFTQGLCAGFLEGGTACQDGGSLVCEDETPEPQLILRG 560

QY 262 VTSWGLGCGRWRNVKRSQDGSFGIFTDISKVLSWIHEH 301
Db 561 IVSWSGCG---NRLK-----PGVYTDVANYLAWIREH 590

RESULT 13
A56318
enteropeptidase (EC 3.4.21.9) precursor - human
N:Alternate names: enterokinase
C:Species: Homo sapiens (man)
C:Date: 19-May-1995 #sequence_revision 09-Aug-1996 #text_change 18-Jun-1999
C:Accession: A56318; B43090
R:Kitamoto, Y.; Veille, R.A.; 1995
Biochemistry 34, 4562-4568, 1995
A:Title: cDNA sequence and chromosomal localization of human enterokinase, the proteo
A:Reference number: A56318; MUID:95234679; PMID:7718557
A:Accession: A56318
A:Molecule type: mRNA
A:Residues: 1-1019 <KIT>
A:Cross-references: GB:U09860; NID:g746412; PIDN:AAC50138.1; PID:g746413
R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease co
A:Reference number: A43090; MUID:94329561; PMID:8052624
A:Accession: B43090
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 749-1019 <KI2>
A:Cross-references: GB:U09860
A:Comment: The mechanism of association with the membrane of the intestinal brush bor
otated below) or with amino-terminal myristoylation of the heavy chain.
C:Genetics:
A:Gene: GDB:PRSS7
A:Cross-references: GDB:384083; OMIM:226200
A:Map position: 21q21-21q21
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and lig
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms inv
ducts.
C:Function:
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
A:Pathway: intestinal digestive hydrolase cascade
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymoge
F:1-784/Product: enteropeptidase heavy chain #status predicted <HCH>
F:22-38/Domain: transmembrane #status predicted <TM>
F:184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:342-504/Domain: MAM homology <MAM>
F:526-631/Domain: C1r/C1s repeat homology <C1R>

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:59:00 ; Search time 26 Seconds
(without alignments)
629.210 Million cell updates/sec

Title: US-09-735-713a-2
Perfect score: 1636
Sequence: 1 MSKMLISRNKLLLLGIVF.....IFTDISKVLWTHIQTGN 306

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues
Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		Published_Applications_AA:*	
1:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*	2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*	4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*	6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*	8:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*	10:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*	12:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*	14:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	100.0	306	9 US-09-735-713a-2	Sequence 2, Appli
2	1618	98.9	302	9 US-09-735-713a-4	Sequence 4, Appli
3	1362	83.3	556	10 US-09-888-615-106	Sequence 106, App
4	898	54.9	164	9 US-09-735-713a-6	Sequence 6, Appli
5	489.5	29.9	802	9 US-09-978-295A-169	Sequence 169, App
6	489.5	29.9	802	9 US-09-978-697-169	Sequence 169, App
7	489.5	29.9	802	9 US-09-978-192A-169	Sequence 169, App
8	489.5	29.9	802	9 US-09-999-832A-169	Sequence 169, App
9	489.5	29.9	802	9 US-09-978-189-169	Sequence 169, App
10	489.5	29.9	802	9 US-09-978-608A-169	Sequence 169, App
11	489.5	29.9	802	9 US-09-978-191A-169	Sequence 169, App
12	489.5	29.9	802	9 US-09-978-403A-169	Sequence 169, App
13	489.5	29.9	802	9 US-09-978-564A-169	Sequence 169, App
14	489.5	29.9	802	9 US-09-978-585A-169	Sequence 169, App
15	489.5	29.9	802	9 US-10-017-081A-169	Sequence 169, App
16	489.5	29.9	802	10 US-09-888-615-113	Sequence 113, App
17	467	28.5	453	9 US-09-978-295A-69	Sequence 69, Appl
18	467	28.5	453	9 US-09-978-697-69	Sequence 69, Appl
19	467	28.5	453	9 US-09-978-192A-69	Sequence 69, Appl

20	467	28.5	453	9 US-09-999-832A-69	Sequence 69, Appl
21	467	28.5	453	9 US-09-978-189-69	Sequence 69, Appl
22	467	28.5	453	9 US-10-174-590-64	Sequence 64, Appl
23	467	28.5	453	9 US-10-176-758-64	Sequence 64, Appl
24	467	28.5	453	9 US-10-175-737-64	Sequence 64, Appl
25	467	28.5	453	9 US-10-173-706-64	Sequence 64, Appl
26	467	28.5	453	9 US-10-175-738-64	Sequence 64, Appl
27	467	28.5	453	9 US-10-175-752-64	Sequence 64, Appl
28	467	28.5	453	9 US-10-176-482-64	Sequence 64, Appl
29	467	28.5	453	9 US-10-176-757-64	Sequence 64, Appl
30	467	28.5	453	9 US-10-176-913-64	Sequence 64, Appl
31	467	28.5	453	9 US-10-180-552-64	Sequence 64, Appl
32	467	28.5	453	9 US-10-180-557-64	Sequence 64, Appl
33	467	28.5	453	9 US-10-173-700-64	Sequence 64, Appl
34	467	28.5	453	9 US-10-174-572-64	Sequence 64, Appl
35	467	28.5	453	9 US-10-174-579-64	Sequence 64, Appl
36	467	28.5	453	9 US-10-174-582-64	Sequence 64, Appl
37	467	28.5	453	9 US-10-174-588-64	Sequence 64, Appl
38	467	28.5	453	9 US-10-175-739-64	Sequence 64, Appl
39	467	28.5	453	9 US-10-175-740-64	Sequence 64, Appl
40	467	28.5	453	9 US-10-175-743-64	Sequence 64, Appl
41	467	28.5	453	9 US-10-176-488-64	Sequence 64, Appl
42	467	28.5	453	9 US-10-176-492-64	Sequence 64, Appl
43	467	28.5	453	9 US-10-176-747-64	Sequence 64, Appl
44	467	28.5	453	9 US-10-176-750-64	Sequence 64, Appl
45	467	28.5	453	9 US-10-176-985-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-09-735-713a-2
; Sequence 2, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713a-2

Query Match 100.0%; Score 1636; DB 9; Length 306;
Best Local Similarity 100.0%; Pred. No. 5.1e-149;
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSLKMILSRNKLILLGIVFFERKSAALSLPKAPSCGQSLVKVQPNYFNIFSRILGGS	60
DB	1	MSLKMILSRNKLILLGIVFFERKSAALSLPKAPSCGQSLVKVQPNYFNIFSRILGGS	60
QY	61	QVEKGSYPWQVSLKQROKHICGGSIVSPQWVITAHAHCIAANRNIVSTLNTVAGEYDLSQTD	120
DB	61	QVEKGSYPWQVSLKQROKHICGGSIVSPQWVITAHAHCIAANRNIVSTLNTVAGEYDLSQTD	120
QY	121	PGEQTLTIETVIIHPHFSTKKPKMDYDIALKMGAFQGHFVGFCICLPRLREOFEAGFC	180
DB	121	PGEQTLTIETVIIHPHFSTKKPKMDYDIALKMGAFQGHFVGFCICLPRLREOFEAGFC	180

```
Qy 181 TTAGWRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Db 181 TTAGWRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Qy 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIFTDISKVLWIEH 300
|||||
Db 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIFTDISKVLWIEH 300
|||||
Qy 301 HIQTGN 306
|||||
Db 301 HIQTGN 306
|||||

RESULT 2
US-09-735-713A-4
; Sequence 4, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1 Human Proteases and
; FILE REFERENCE: Lex-0108-USA
; CURRENT APPLICATION NUMBER: US/09735,713A
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713A-4

Query Match 98.9%; Score 1618; DB 9; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.7e-147; Indels 0; Gaps 0;
Matches 302; Conservative 0; Mismatches 0;

Qy 5 MLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGSGVEK 64
|||||
Db 1 MLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGSGVEK 60
|||||
Qy 65 GSYPMQVSLKQKQKHICGGSIVSPQWVITAHCIAANRNIVSTLNTVAGEYDLSQTDPEQ 124
|||||
Db 61 GSYPMQVSLKQKQKHICGGSIVSPQWVITAHCIAANRNIVSTLNTVAGEYDLSQTDPEQ 120
|||||
Qy 125 TLTIETVLIHPHFSTKPKMDYDIALLMKAGAFQGFHFGVPGICLPDLREQFAGFICTTAG 184
|||||
Db 121 TLTIETVLIHPHFSTKPKMDYDIALLMKAGAFQGFHFGVPGICLPDLREQFAGFICTTAG 180
|||||
Qy 195 WGRTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDACQDGS 244
|||||
Db 181 WGRTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDACQDGS 240
|||||
Qy 245 GGSMLCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIFTDISKVLWIEH 304
|||||
Db 241 GGSMLCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIFTDISKVLWIEH 300
|||||
Qy 305 GN 306
||
Db 301 GN 302
||

RESULT 3
US-09-888-615-106
; Sequence 106, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
```

```
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-106

Query Match 83.3%; Score 1362; DB 10; Length 556;
Best Local Similarity 85.7%; Pred. No. 2e-122; Indels 8; Gaps 2;
Matches 263; Conservative 8; Mismatches 28;

Qy 1 MSLKMLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGSGS 60
|||||
Db 1 MSLKMLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGSGS 60
|||||
Qy 61 QVEKGSYPWQVSLKQKQKHICGGSIVSPQWVITAHCIAANRNIVSTLNTVAGEYDLSQTD 120
|||||
Db 61 QVEKGSYPWQVSLKQKQKHICGGSIVSPQWVITAHCIAANRNIVSTLNTVAGEYDLSQTD 120
|||||
Qy 121 PGEOTLTETVLIHPHFSTKPKMDYDIALLMKAGAFQGFHFGVPGICLPDLREQFAGFIC 180
|||||
Db 121 PGEOTLTETVLIHPHFSTKPKMDYDIALLMKAGAFQGFHFGVPGICLPDLREQFAGFIC 180
|||||
Qy 181 TTAGWRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Db 181 TTAGWRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Qy 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGP-----GIFTDISK 293
|||||
Db 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGP-----GIFTDISK 293
|||||
Qy 294 VLSWIEH 300
||
Db 300 CGTWCE 306
||

RESULT 4
US-09-735-713A-6
; Sequence 6, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1 Human Proteases and
; FILE REFERENCE: Lex-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
```


US-09-735-713A-6

Query Match 54.9%; Score 898; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 MDYDIALKMGAFQGFHFVGPICLPRLREQFEAGFICTAGWRLTGGVLSQVLQEVN 202
|||||
Db 1 MDYDIALKMGAFQGFHFVGPICLPRLREQFEAGFICTAGWRLTGGVLSQVLQEVN 60

QY 203 LPILTWEECVAALLTKRPISGKTKTLCTGFPDGGDRDACQDGGSGSLMCRNKKGANTLAGV 262
|||||
Db 61 LPILTWEECVAALLTKRPISGKTKTLCTGFPDGGDRDACQDGGSGSLMCRNKKGANTLAGV 120

QY 263 TSWGLGCGGRWNNVRKSDQSGPGIFTDISKVLNHEHIQTGN 306
|||||
Db 121 TSWGLGCGGRWNNVRKSDQSGPGIFTDISKVLNHEHIQTGN 164

RESULT 5

US-09-978-295A-169
; Sequence 169, Application US/09978295A
; Patent No. US20020156006A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2630P1C11

; CURRENT APPLICATION NUMBER: US/09/978,295A

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/064249

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077641

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21

```

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 29.9%; Score 489.5; DB 9; Length 802;
Best Local Similarity 39.3%; Pred. No. 1.4e-38;
Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps

QY 54 SRLGGQVEKSGYPMQVSLKQRKHICGSGIVSPQWVITAAHCIAANRNVSTL--NVTA 111
Db 566 SRVIGVASSEGEWPMQASLQVRGRHICGGALLIADRWITAAHCFQEDSMASVLTWTVL 625
QY 112 GEYDLSTQDGEOTLIETVIIHPHSTKKPMDYDIALKMKAGAFQGFHVGPICLPELR 171
Db 626 GKYVNSRWPFGEVSFKVSRLLHP-YHEEDSHDYDALLQDHPVVRSAARVVCULPARS 684
QY 172 EQFEAGFICTAGWGRTGEGVLSQVLQEVNLPILTWECCVAALLTLKRPISGKTFCLCTG 231
Db 685 HFEPGLCHWITGVALREGGPSINALQKVDVQLIPQDLCSSEAYRQVTP-----RMLCAG 740
QY 232 FPDGGRDACOGDGGSLMCRNKKGAWTLAGVTSWGLGCGRGMNVRKNSDQGSP----GIF 288
Db 741 YRKGGKDACOGDGGGLPVCKALSGRFLAGLVSWGLGCGR-----PNYFGVY 787
QY 289 TDISKVLSWIHE 300
Db 788 TRITGVISWIOQ 799

RESULT 6
US-09-978-697-169
; Sequence 169, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13

```

;	PRIOR APPLICATION NUMBER: 60/066364	;
;	PRIOR FILING DATE: 1997-11-21	;
;	PRIOR APPLICATION NUMBER: 60/077450	;
;	PRIOR FILING DATE: 1998-03-10	;
;	PRIOR APPLICATION NUMBER: 60/077632	;
;	PRIOR FILING DATE: 1998-03-11	;
;	PRIOR APPLICATION NUMBER: 60/077641	;
;	PRIOR FILING DATE: 1998-03-11	;
;	PRIOR APPLICATION NUMBER: 60/077649	;
;	PRIOR FILING DATE: 1998-03-11	;
;	PRIOR APPLICATION NUMBER: 60/077791	;
;	PRIOR FILING DATE: 1998-03-12	;
;	PRIOR APPLICATION NUMBER: 60/078004	;
;	PRIOR FILING DATE: 1998-03-13	;
;	PRIOR APPLICATION NUMBER: 60/078886	;
;	PRIOR FILING DATE: 1998-03-20	;
;	PRIOR APPLICATION NUMBER: 60/078936	;
;	PRIOR FILING DATE: 1998-03-20	;
;	PRIOR APPLICATION NUMBER: 60/078910	;
;	PRIOR FILING DATE: 1998-03-20	;
;	PRIOR APPLICATION NUMBER: 60/078939	;
;	PRIOR FILING DATE: 1998-03-20	;
;	PRIOR APPLICATION NUMBER: 60/079294	;
;	PRIOR FILING DATE: 1998-03-25	;
;	PRIOR APPLICATION NUMBER: 60/079656	;
;	PRIOR FILING DATE: 1998-03-26	;
;	PRIOR APPLICATION NUMBER: 60/079664	;
;	PRIOR FILING DATE: 1998-03-27	;
;	PRIOR APPLICATION NUMBER: 60/079689	;
;	PRIOR FILING DATE: 1998-03-27	;
;	PRIOR APPLICATION NUMBER: 60/079663	;
;	PRIOR FILING DATE: 1998-03-27	;
;	PRIOR APPLICATION NUMBER: 60/079728	;
;	PRIOR FILING DATE: 1998-03-27	;
;	PRIOR APPLICATION NUMBER: 60/079786	;
;	PRIOR FILING DATE: 1998-03-27	;
;	PRIOR APPLICATION NUMBER: 60/079920	;
;	PRIOR FILING DATE: 1998-03-30	;
;	PRIOR APPLICATION NUMBER: 60/079923	;
;	PRIOR FILING DATE: 1998-03-30	;
;	PRIOR APPLICATION NUMBER: 60/080105	;
;	PRIOR FILING DATE: 1998-03-31	;
;	PRIOR APPLICATION NUMBER: 60/080107	;
;	PRIOR FILING DATE: 1998-03-31	;
;	PRIOR APPLICATION NUMBER: 60/080165	;
;	PRIOR FILING DATE: 1998-03-31	;
;	PRIOR APPLICATION NUMBER: 60/080194	;
;	PRIOR FILING DATE: 1998-03-31	;
;	PRIOR APPLICATION NUMBER: 60/080327	;
;	PRIOR FILING DATE: 1998-04-01	;
;	PRIOR APPLICATION NUMBER: 60/080328	;
;	PRIOR FILING DATE: 1998-04-01	;
;	PRIOR APPLICATION NUMBER: 60/080333	;
;	PRIOR FILING DATE: 1998-04-01	;
;	PRIOR APPLICATION NUMBER: 60/080334	;
;	PRIOR FILING DATE: 1998-04-01	;
;	PRIOR APPLICATION NUMBER: 60/081070	;
;	PRIOR FILING DATE: 1998-04-08	;
;	PRIOR APPLICATION NUMBER: 60/081049	;
;	PRIOR FILING DATE: 1998-04-08	;
;	PRIOR APPLICATION NUMBER: 60/081203	;
;	PRIOR FILING DATE: 1998-04-09	;
;	PRIOR APPLICATION NUMBER: 60/081229	;
;	PRIOR FILING DATE: 1998-04-09	;
;	PRIOR APPLICATION NUMBER: 60/081955	;
;	PRIOR FILING DATE: 1998-04-15	;
;	PRIOR APPLICATION NUMBER: 60/081817	;
;	PRIOR FILING DATE: 1998-04-15	;
;	PRIOR APPLICATION NUMBER: 60/081819	;

1	PRIOR FILING DATE: 1998-04-15
2	PRIOR APPLICATION NUMBER: 60/081952
3	PRIOR FILING DATE: 1998-04-15
4	PRIOR APPLICATION NUMBER: 60/081838
5	PRIOR FILING DATE: 1998-04-15
6	PRIOR APPLICATION NUMBER: 60/082568
7	PRIOR FILING DATE: 1998-04-21
8	PRIOR APPLICATION NUMBER: 60/082569
9	PRIOR FILING DATE: 1998-04-21
10	PRIOR APPLICATION NUMBER: 60/082704
11	PRIOR FILING DATE: 1998-04-22
12	PRIOR APPLICATION NUMBER: 60/082804
13	PRIOR FILING DATE: 1998-04-22
14	PRIOR APPLICATION NUMBER: 60/082700
15	PRIOR FILING DATE: 1998-04-22
16	PRIOR APPLICATION NUMBER: 60/082797
17	PRIOR FILING DATE: 1998-04-22
18	PRIOR APPLICATION NUMBER: 60/082796
19	PRIOR FILING DATE: 1998-04-23
20	PRIOR APPLICATION NUMBER: 60/083336
21	PRIOR FILING DATE: 1998-04-27
22	PRIOR APPLICATION NUMBER: 60/083322
23	PRIOR FILING DATE: 1998-04-28
24	PRIOR APPLICATION NUMBER: 60/083392
25	PRIOR FILING DATE: 1998-04-29
26	PRIOR APPLICATION NUMBER: 60/083495
27	PRIOR FILING DATE: 1998-04-29
28	PRIOR APPLICATION NUMBER: 60/083496
29	PRIOR FILING DATE: 1998-04-29
30	PRIOR APPLICATION NUMBER: 60/083499
31	PRIOR FILING DATE: 1998-04-29
32	PRIOR APPLICATION NUMBER: 60/083545
33	PRIOR FILING DATE: 1998-04-29
34	PRIOR APPLICATION NUMBER: 60/083554
35	PRIOR FILING DATE: 1998-04-29
36	PRIOR APPLICATION NUMBER: 60/083558
37	PRIOR FILING DATE: 1998-04-29
38	PRIOR APPLICATION NUMBER: 60/083559
39	PRIOR FILING DATE: 1998-04-29
40	PRIOR APPLICATION NUMBER: 60/083500
41	PRIOR FILING DATE: 1998-04-29
42	PRIOR APPLICATION NUMBER: 60/083742
43	PRIOR FILING DATE: 1998-04-30
44	PRIOR APPLICATION NUMBER: 60/084366
45	PRIOR FILING DATE: 1998-05-05
46	PRIOR APPLICATION NUMBER: 60/084414
47	PRIOR FILING DATE: 1998-05-06
48	PRIOR APPLICATION NUMBER: 60/084441
49	PRIOR FILING DATE: 1998-05-06
50	PRIOR APPLICATION NUMBER: 60/084637
51	PRIOR FILING DATE: 1998-05-07
52	PRIOR APPLICATION NUMBER: 60/084639
53	PRIOR FILING DATE: 1998-05-07
54	PRIOR APPLICATION NUMBER: 60/084640
55	PRIOR FILING DATE: 1998-05-07
56	PRIOR APPLICATION NUMBER: 60/084598
57	PRIOR FILING DATE: 1998-05-07
58	PRIOR APPLICATION NUMBER: 60/084600
59	PRIOR FILING DATE: 1998-5-07
60	PRIOR APPLICATION NUMBER: 60/084627
61	PRIOR FILING DATE: 1998-05-07
62	PRIOR APPLICATION NUMBER: 60/084643
63	PRIOR FILING DATE: 1998-05-07
64	PRIOR APPLICATION NUMBER: 60/085339
65	PRIOR FILING DATE: 1998-05-13
66	PRIOR APPLICATION NUMBER: 60/085338
67	PRIOR FILING DATE: 1998-05-13
68	PRIOR APPLICATION NUMBER: 60/085323
69	PRIOR FILING DATE: 1998-05-13
70	PRIOR APPLICATION NUMBER: 60/085582
71	PRIOR FILING DATE: 1998-05-15
72	PRIOR APPLICATION NUMBER: 60/085700
73	PRIOR FILING DATE: 1998-05-15

[illegible]

APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1G63
CURRENT APPLICATION NUMBER: US/09/999,832A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07


```
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC22
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 169
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-608A-169

Query Match          29.9%; Score 489.5; DB 9; Length 802;
Best Local Similarity 39.3%; Pred. No. 1.4e-38;
Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps 5;

QY 54 SRILGSGVKEGSPQVSLKQKQKHCIGGSIVSPQWVITAAHCIAANRIVSTL--NNTA 111
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 566 SRIVGASVSSEGEWPQASLQVRGRHICGALIDRWVITAAHCFQEDSMASTVLTWFL 625
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 112 GEYDLSQTDPGQTLTIETVIITHPEFTKKPMDYDIALKMGAGQFGHFVGPICLPRL 171
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 626 GRVQNSRNPGEVSEKSKLLLHP-YHEEDSHDYDVALQLDHPVVRSAARVPCLPARS 684
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 172 EQFEAGFICTTAGWRLTEGGVLSQVLQVNLPILTWEECVAAALLTLKRPISGKTFCLCTG 231
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 685 HFEFGLHCWITGWLALREGGPISNALQKVDVQLIPQDLCSAYRYQVTP----RMLCAG 740
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 232 FPDGGRDACQDGSGLMCRNKKGAWTLAGVTSWGLGCGGRNVRNKSQDQSP---GIF 288
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 741 YRKGKDACQDGSGLGKALSGRWFLAGLSWGLGCGR-----PNYFGVY 787
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 289 TDISKVLSIHE 300
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 788 TRITGVISWIIQ 799
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 11
US-09-798-191A-169
; Sequence 169, Application US/09978191A
; Publication No. US20030050239A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
```

```
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC4
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
```

RESULT 12
US-09-978-403A-169
; Sequence 169, Application US/09978403A

; Publication No. US20030050240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Oiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Kijavlin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2630P1C17
 ; CURRENT APPLICATION NUMBER: US/09/978,403A
 ; CURRENT FILING DATE: 2002-03-19
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079664
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079663

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:57:25 ; Search time 22 Seconds
(without alignments)
409.246 Million cell updates/sec

Title: US-09-735-713A-2

Perfect score: 1636

Sequence: 1 MSKMLISRNKLLILLGIVF.....IFTDISKVLWIHIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	467	28.5	454	US-09-518-046-2	Sequence 2, Appli
2	460	28.1	248	US-08-944-483-63	Sequence 63, Appli
3	447.5	27.4	638	US-08-681-151-3	Sequence 3, Appli
4	446.5	27.3	455	US-09-261-416-2	Sequence 2, Appli
5	444.5	27.2	328	US-09-386-642-11	Sequence 11, Appli
6	442.5	27.0	798	US-08-200-900A-2	Sequence 2, Appli
7	442.5	27.0	798	US-08-200-900A-2	Sequence 2, Appli
8	440.5	26.9	416	US-09-000-846-2	Sequence 12, Appli
9	438	26.8	319	US-09-386-642-12	Sequence 3, Appli
10	437.5	26.7	256	US-09-027-337-3	Sequence 3, Appli
11	437.5	26.7	256	US-09-644-600-3	Sequence 3, Appli
12	432.5	26.4	255	US-08-944-483-67	Sequence 67, Appli
13	432.5	26.4	400	US-09-004-731-30	Sequence 30, Appli
14	432.5	26.4	400	US-09-004-731-33	Sequence 33, Appli
15	432.5	26.4	400	US-08-749-699-30	Sequence 30, Appli
16	432.5	26.4	400	US-08-749-699-33	Sequence 33, Appli
17	432.5	26.4	400	US-09-004-729-30	Sequence 30, Appli
18	432.5	26.4	400	US-09-004-729-33	Sequence 33, Appli
19	430.5	26.3	387	US-09-032-215-8	Sequence 8, Appli
20	430.5	26.3	387	US-09-032-215-13	Sequence 13, Appli
21	430	26.3	299	US-08-944-483-66	Sequence 66, Appli
22	430	26.3	315	US-09-386-653A-9	Sequence 9, Appli
23	429	26.2	290	US-09-386-653A-7	Sequence 7, Appli
24	426	26.0	242	US-09-004-731-36	Sequence 36, Appli
25	426	26.0	242	US-08-749-699-36	Sequence 36, Appli
26	426	26.0	242	US-09-004-729-36	Sequence 36, Appli
27	425.5	26.0	235	US-08-807-151-3	Sequence 3, Appli

28	425.5	26.0	235	4	US-09-478-957-3	Sequence 3, Appli
29	421.5	25.8	317	4	US-09-386-629-7	Sequence 7, Appli
30	419.5	25.6	235	4	US-08-944-483-65	Sequence 65, Appli
31	419.5	25.6	237	4	US-08-163-919A-3	Sequence 3, Appli
32	419.5	25.6	237	5	US-09-386-629-8	Sequence 8, Appli
33	418	25.6	327	4	US-09-008-271A-6	Sequence 2, Appli
34	414	25.3	435	4	US-09-656-002-2	Sequence 6, Appli
35	413	25.2	423	4	US-08-944-483-64	Sequence 64, Appli
36	408	24.9	238	3	US-08-807-151-1	Sequence 1, Appli
37	408	24.9	283	4	US-09-478-957-1	Sequence 1, Appli
38	407	24.9	492	4	US-09-342-749-2	Sequence 2, Appli
39	407	24.9	492	4	US-09-691-840-2	Sequence 2, Appli
40	406	24.8	254	4	US-09-578-303-5	Sequence 5, Appli
41	405	24.8	231	2	US-09-027-337-6	Sequence 6, Appli
42	405	24.8	231	4	US-09-644-600-6	Sequence 6, Appli
43	405	24.8	231	4	US-09-008-271A-3	Sequence 3, Appli
44	403	24.6	314	4	US-09-016-366A-23	Sequence 23, Appli
45	402	24.6	267	2	US-09-016-366A-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1

US-09-518-046-2

; Sequence 2, Application US/09518046

; Patent No. 6294663

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed

; FILE REFERENCE: D6192CIP

; CURRENT APPLICATION NUMBER: US/09/518,046

; EARLIER FILING DATE: 2000-03-02

; EARLIER FILING DATE: 1999-03-03

; NUMBER OF SEQ ID NOS: 153

; SEQ ID NO 2

; LENGTH: 454

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: complete amino acid sequence of TADG-12

; OTHER INFORMATION: protein

US-09-518-046-2

Query Match 28.5%; Score 467; DB 4; Length 454;
Best Local Similarity 32.1%; Pred. No. 4.5e-42;
Matches 99; Conservative 61; Mismatches 114; Indels 34; Gaps 9;

QY	1	MSKMLISRNKLLILLGIVFFERG-KSAALSLPKAPSCGSLVKVQPNVFNFSRLGG	59
DB	171	VSIDHLLPDKVTALHHSYVREGSCASGHVVTLLQCTACGHRH-----GYSSRVGG	221
QY	60	SOVEKSGYPQVSLKQKHICGSGIVSPQWVITAAHCIAANRNIVSTLNTVAGEYDLSQT	119
DB	222	NMSLLSQWPQASLQFGYHLCGSGVITPLWITAAHCIVDLYLPKSWTQVGLVSLD-	280
QY	120	DPEQTITITVTHFSTKKPMYDIALKMGAFQFGHFGVGPICLPRLREQEAGFI	179
DB	281	NPAPSHL-VEKIVYHSYKPKR-LGNDIALMKLAGLTFNEMIQPVCLPSEENFPDGKV	338
QY	180	CTTAGWRLTEGGVLSVQLEVNLPILTWEBCVAALLTLKRPIG-----KTFICTGPPDG	235
DB	339	CWTSGWGTEDGGDASPLNHAAPLISNKC-----NHRDVYGGIISPMCLCAGLYTG	392
QY	236	GRDACQDGSGLMCRNKKGAWTLAGVTSNCLGCGRNRNRVRSKDSGSGIFTDISKVL	295
DB	393	GVDSQCQDGSGLPVCQBRR-LWKLVGATSFSGIC-----AEVNKPGVTRVTSFL	441
QY	296	SWTHEIQ 303	

10

```
QY 194 LSOVLQEVNLPILTWEECVAAALLTLKRPISGKFTICTGFPDGRDACQDGGSGSLMCRNK 253
      :||: :||: ||| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 531 TQNILKATIPLVNEECQK---KYRDYVITKQMICAGYKEGGIDACKDGGSGPLVCKH- 586

QY 254 KGAWTLAGVTSWGLCGGRWRNNVRKSDGSGPQIFTDISKVLSWIHEHIQT 304
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 587 SGRWQLVGTISWEGGAR-----KEQPGVYTKVAEYIDWILEKIQS 627

RESULT 4:
US-09-261-416-2
; Sequence 2, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TAGD-12: A No. 6291663a1 Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of TAGD-12 encoded by nucleotides
; OTHER INFORMATION: 144 to 1511 of Sequence 1
; Patent No. 6291663
US-09-261-416-2

Query Match 27.3%; Score 446.5; DB 4; Length 455;
Best Local Similarity 31.0%; Pred. No. 7.5e-40;
Matches 98; Conservative 57; Mismatches 112; Indels 49; Gaps 10;

QY 1 MSLKMLISRNKLLILGIVFFERG-KSAALSLPKAPSCGQSLVKVQPNVNFISRLGG 59
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 171 VSDHLLPDDKVTALHSHVYVREGCAGHVVLQCTACGHR-----GYSSRIVGG 221

QY 60 SQVEKSGYPWQVSLKQKQKHICGGSVSPQWVITAAHCIAANRNIVSTLNVTAGEYDLSQT 119
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 222 NMSLLSQWPQASQFQYHLCGSGVITPLWITTAHCYVDLYLPKSWITQVGLVSLD- 280

QY 120 DPGEQTLTETVTHPHFSTKPMQDYDIALKMGAFQFHFVGPICLPRLPEQFEAGFI 179
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 281 NPAPSHL-VEKIVYHSHKPKR-LGNDIALMKLAGPLTFNEMIQPVCLPNSSENFDPGKV 338

QY 180 CTTAGHRLTEGGVLSQVLEQVNLPLT-----WEECVAAALLTLKRPISGKTFL 228
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 339 CWTSGHGATEDGGDASFLNHAAPVPLISKNKDLOPQGRVRHHLLPLHAL----- 386

QY 229 CTGFPDGR-DACQDGGSGSLMCRNKGAWTLAGVTSWGLCGGRWRNNVRKSDGSGPGI 287
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 387 -RGLPDGWRNNSCGDGGPLVCQERR-LWKLVGATSFGIGC-----ADVKNPGV 434

QY 288 FTDISKVLSWIHEHIQ 303
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 435 YTRVTSFLDWIHEQME 450

RESULT 5
US-09-386-642-11
; Sequence 11, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
```

```
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-11

Query Match 27.2%; Score 444.5; DB 4; Length 328;
Best Local Similarity 32.3%; Pred. No. 7.8e-40;
Matches 100; Conservative 52; Mismatches 119; Indels 39; Gaps 8;

QY 3 LKMLISRNKLLILGIVFFERG-----KSAALSLPKAPSCGQSLVKVQPNVNFISRLG 58
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 12 LULLLVVSNLLCQGVVSDYKDDDDVDAAALAAP-----FDDDDKIVG 54

QY 59 GSOVEKSGYPWQVSLKQKQKHICGGSVSPQWVITAAHCIAANRNIVSTLNVTAGEYDLSQ 118
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 55 GYALEAGQWPQVSYTYEGVHVCQSLVSEQWVLSAAHCFPSHHEAVEYKLGALHQLDS 114

QY 119 TDPGEQTLTETVTHPHFSTKPMQDYDIALKMGAFQFHFVGPICLPRLPEQFEAGF 178
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 115 YSEDAKYSTLKDIIPHSY-LQEGSQGDIALQLSRPITFSYRIRPICLPAANASFPNGL 173

QY 179 ICTTAGWRLTEGGVL--SQVLEQVNLPLTWEECVAAALLTLKRP-----ISGKTFICTGF 232
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 174 HCTVTGWGHVAPSVSLTPKPLQQLLEVPLISRETCLYNIDAKPEEPHFVQEDMVCAGY 233

QY 233 PDGGRDACQDGGSGSLMCRNKGAWTLAGVTSWGLCGGRWRNNVRKSDGSGPIETDIS 292
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 234 VEGKDACQDGGSGPLSC-PVEGLWYLTGIVSWGACGA--RNR-----PGVYTLAS 282

QY 293 KYLSWIHEHI 302
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 283 SYASWISQSKV 292

RESULT 6
US-08-200-900A-2
; Sequence 2, Application US/08200900A
; Patent No. 5665566
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc. - Legal Affairs
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/200,900A
; FILING DATE: 23-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meinert, Maureen C.
; REGISTRATION NUMBER: 31,544
; REFERENCE/DOCKET NUMBER: GI 5201-FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170 X8574
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-200-900A-2

Query Match          27.0%; Score 442.5; DB 1; Length 798;
Best Local Similarity 34.9%; Pred. No. 4.5e-39;
Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;

QY 36 SCGSLV--KVQPNYNIIFSRILGGSGVEKGYSPWVSLKQKHKICGGSIYSPQWVIT 93
Db 550 SCGKLLVQEVSP-----KIVGSDSREGAPWVVALYFDQVCGASLSRDLWLS 601
QY 94 AAHCIAIRNI-VSTLNTAGYDLSQ-TDPEQTLTIETVLIHHPFSTKKRPMYDIALLK 151
Db 602 AAHCYGRNMEPSKWKAVLGLHMASNLTSPOIETRLIDQIVINPHYN-KRRKNDIAMMH 660
QY 152 MAGAFQFCHFGVPCICLPRLRQFEAGFICTAGWGRLTGEGVLSQVLEQVNLPLTLTWEEC 211
Db 661 LEMKVNTDYIQICLPLEENQVPPGRICSTAGWALYQGSTADVLQEADVPLLSNEKC 720
QY 212 VAAL----LTLKRPISGKTFCTGFPDGGDRDACQDGGSLMCRNKKGAWTLAGVTSNGL 267
Db 721 QQQMPYNI-----ENMVACGYEAGGVDSQDGGSGPLMCQ-ENNRWLLAGVTSFGY 772
QY 268 GCGRGNRNVRKSDQSGPGIFTDLSKVLSTIHEHI 302
Db 773 QCALPNR-----PGVYARVPRETETIQSFL 797

RESULT 7
PCT-US94-00616-2
; Sequence 2, Application PC/TUS9400616
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF SEQUENCES: 33
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00616
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-00616-2

Query Match          27.0%; Score 442.5; DB 5; Length 798;
Best Local Similarity 34.9%; Pred. No. 4.5e-39;
Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;

QY 36 SCGSLV--KVQPNYNIIFSRILGGSGVEKGYSPWVSLKQKHKICGGSIYSPQWVIT 93
Db 550 SCGKLLVQEVSP-----KIVGSDSREGAPWVVALYFDQVCGASLSRDLWLS 601
QY 94 AAHCIAIRNI-VSTLNTAGYDLSQ-TDPEQTLTIETVLIHHPFSTKKRPMYDIALLK 151
Db 602 AAHCYGRNMEPSKWKAVLGLHMASNLTSPOIETRLIDQIVINPHYN-KRRKNDIAMMH 660
QY 152 MAGAFQFCHFGVPCICLPRLRQFEAGFICTAGWGRLTGEGVLSQVLEQVNLPLTLTWEEC 211
Db 661 LEMKVNTDYIQICLPLEENQVPPGRICSTAGWALYQGSTADVLQEADVPLLSNEKC 720
```

```
QY 212 VAAL----LTLKRPISGKTFCTGFPDGGDRDACQDGGSLMCRNKKGAWTLAGVTSNGL 267
Db 721 QQQMPYNI-----ENMVACGYEAGGVDSQDGGSGPLMCQ-ENNRWLLAGVTSFGY 772
QY 268 GCGRGNRNVRKSDQSGPGIFTDLSKVLSTIHEHI 302
Db 773 QCALPNR-----PGVYARVPRETETIQSFL 797

RESULT 8
US-09-000-846-2
; Sequence 2, Application US/09000846
; Patent No. 5981830
; GENERAL INFORMATION:
; APPLICANT: WU, QINGYU
; APPLICANT: SADLER, JASPER
; TITLE OF INVENTION: KNOCKOUT MICE AND THEIR PROGENY WITH
; TITLE OF INVENTION: DISRUPTED SERINE PROTEASE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,846
; FILING DATE: 30-DEC-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/866,058
; FILING DATE: 30-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: LEBOVITZ, RICHARD M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: BERLX 65P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 416 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-000-846-2

Query Match          26.9%; Score 440.5; DB 2; Length 416;
Best Local Similarity 33.0%; Pred. No. 2.9e-39;
Matches 102; Conservative 48; Mismatches 112; Indels 47; Gaps 11;

QY 11 KLILLGIVFERGKSAALSAPRACSCQSLVKVQPNYNIIFSRILGGSGVEKGYSPWQ 70
Db 129 RLLDVISVDCPRGFLTAT---CQDGRRLPV-----DRIVGGQDSSLGRNPWQ 176
QY 71 VSLKQKHKICGGSIYSPQWVITAAHCIAIRNI-TVSTLNTAGYDLSQDTPDGTLTIE 129
Db 177 VSURYDTHLCGSLSGDWLTAAHCFPBRNRVLSRNRVFAGA--VARTSPHAVQLGVQ 234
QY 130 TVIHPHF-----STKKPMYDIALKKAGAFQFHEVGPICLPRLRQFEAGFICTAG 184
Db 235 AVIYHGYLPFROPTIDENSNDIALVHLSLPLTEYIQVCLPAAQALVDGKVCITVG 294
QY 185 WGRITGEGVLSQVLEQVNLPLTLTWEECAALLTLKRP-----ISGKTFCTGFPDGGDR 238
Db 295 WGNTOFYGOAMVLEQARVPLISNEVC-----NSPDFYGNQIKPKMF-CAGYPEGGID 346
```

Qy	239	ACOGDGGSLMCRNK---	KGAWTLAGVTSWGLCGGRGNRRNNVRKSSDGSGPGIFTLSKVL	295	
		:	: :		
Db	347	ACOGDGGPPGEDSISGT	SRNLGCIVSWGTCAL----	ARK-----PGVYTKVTDFR	396
Qy	296	SWIHEHIQT	304		
		:	:		
Db	397	EWIFKAIKT	405		

```

RESULT 9
US-09-386-642-12
; Sequence 12, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-12

```

RESULT 10
US-09-027-337-3
; Sequence 3, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
; TITLE OF INVENTION: Breast and Ovarian Carcinomas
; FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027.337B

[illegible]

```

RESULT 11
US-09-644-600-3
; Sequence 3, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirocoshi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644.600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 3
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Hepsin
US-09-644-600-3

```


Qy 266 GLCGRGWRNNVRKSDQSPGIFTDISKVLWSI 298
|:| | | : | | : | | : | | : | |
Db 373 GIGGR-----EHLPGVVTNIQKFIPIWI 395

Search completed: March 25, 2003, 04:03:36
Job time : 28 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 25, 2003, 01:18:53 ; Search time 59 Seconds
(without alignments)
1590.561 Million cell updates/sec

Title: US-09-735-713a-2

Perfect score: 1636

Sequence: 1 MSKMLISRNKLLIGIVF.....IFTDISKVLWIHIQTGN 306

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09735713/runat_18032003_124142_21610/app_query.fasta.1.455
-DB=issued_Patents_NA -QFWT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713.ecgn.1.1.40.0runat_18032003_124142_21610 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	462	28.2	2413	4	US-09-518-046-1
2	446.5	27.3	2416	4	US-09-261-416-1
3	444.5	27.2	1189	4	US-09-386-642-7
4	444.5	27.2	1783	4	US-09-510-738A-188
5	444.5	27.2	2363	4	US-09-742-703-3
6	442.5	27.0	2581	1	US-08-200-900A-1
7	442.5	27.0	2581	5	PCT-US94-00616-1
8	440.5	26.9	1605	2	US-09-000-846-1
9	438	26.8	1142	4	US-09-386-642-8
10	435	26.6	2544	4	US-09-518-046-3
11	434.5	26.6	1200	4	US-09-004-731-32
12	434.5	26.6	1200	4	US-09-004-731-34

13	434.5	26.6	1200	4	US-08-749-699-32	Sequence 32, Appl
14	434.5	26.6	1200	4	US-08-749-699-34	Sequence 34, Appl
15	434.5	26.6	1200	4	US-09-004-729-32	Sequence 32, Appl
16	434.5	26.6	1200	4	US-09-004-729-34	Sequence 34, Appl
17	434.5	26.6	1894	4	US-09-004-731-29	Sequence 29, Appl
18	434.5	26.6	1894	4	US-09-004-731-31	Sequence 31, Appl
19	434.5	26.6	1894	4	US-09-032-215-3	Sequence 3, Appl
20	434.5	26.6	1894	4	US-08-749-699-29	Sequence 29, Appl
21	434.5	26.6	1894	4	US-08-749-699-31	Sequence 31, Appl
22	434.5	26.6	1894	4	US-09-004-729-29	Sequence 29, Appl
23	434.5	26.6	1894	4	US-09-004-729-31	Sequence 31, Appl
24	432.5	26.4	1161	4	US-09-032-215-9	Sequence 9, Appl
25	432.5	26.4	1161	4	US-09-032-215-10	Sequence 10, Appl
26	432.5	26.4	1161	4	US-09-032-215-15	Sequence 15, Appl
27	432.5	26.4	1161	4	US-09-032-215-16	Sequence 16, Appl
28	432.5	26.4	1175	4	US-09-032-215-12	Sequence 12, Appl
29	432.5	26.4	1175	4	US-09-032-215-14	Sequence 14, Appl
30	432.5	26.4	1373	4	US-09-032-215-6	Sequence 6, Appl
31	432.5	26.4	1373	4	US-09-032-215-7	Sequence 7, Appl
32	430	26.3	1130	4	US-09-386-653A-8	Sequence 8, Appl
33	429	26.2	1110	4	US-09-386-653A-1	Sequence 1, Appl
34	426	26.0	726	4	US-09-004-731-35	Sequence 35, Appl
35	426	26.0	726	4	US-08-749-699-35	Sequence 35, Appl
36	426	26.0	726	4	US-09-004-729-35	Sequence 35, Appl
37	421.5	25.8	1430	4	US-09-386-629-1	Sequence 1, Appl
38	418	25.6	1166	4	US-09-386-629-2	Sequence 2, Appl
39	414	25.3	2038	4	US-09-008-271A-18	Sequence 18, Appl
40	413	25.2	2079	4	US-09-656-002-1	Sequence 1, Appl
41	408	24.9	1077	3	US-08-807-151-2	Sequence 2, Appl
42	408	24.9	1077	4	US-09-478-957-2	Sequence 2, Appl
43	407	24.9	1479	4	US-09-342-749-1	Sequence 1, Appl
44	407	24.9	1479	4	US-09-691-840-1	Sequence 1, Appl
45	403	24.6	1081	4	US-09-008-271A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-518-046-1
; Sequence 1, Application US/09518046
; Patent No. 6294663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; FILE OF INVENTION: in Ovarian Carcinoma and Uses Thereof
; FILE REFERENCE: D6192CIP
; CURRENT APPLICATION NUMBER: US/09/518,046
; CURRENT FILING DATE: 2000-03-02
; EARLIER APPLICATION NUMBER: 09/261,416
; EARLIER FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 1
; LENGTH: 2413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; OTHER INFORMATION: entire cDNA sequence of TADG-12 gene
US-09-518-046-1

Alignment Scores:
Pred. No.: 3.08e-42 Length: 2413
Score: 462.00 Matches: 98
Percent Similarity: 51.95% Conservative: 62
Best Local Similarity: 31.82% Mismatches: 114
Query Match: 28.24% Indels: 34
DB: 4 Gaps: 9

US-09-735-713A-2 (1-306) x US-09-518-046-1 (1-2413)
QY 1 MetSerLeuLysMetLeuLysSerArgAsnLysLeuLeuLeuLeuGlyIleValPhe 20
:::||||::: |||::: :|||::: |||
|||:::

```
Db 654 GTCTCATCATCACTCTTGCAGATGACAAAGGTGACTGCAATTACCACTCAGTATAT 713
QY 21 PheGluArgGly---LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGln 39
Db 714 GTGAGGAGGAGTGCCTCTGCCACGTGGTTACCTTGAGTGCACACGCTGTGTCAT 773
QY 40 SerLeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGly 59
Db 774 AGAAGG-----GGCTACAGCTCAGCATCGTGGGTGA 806
QY 60 SerGlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHis 79
Db 807 AACATGCTCCTGCTCTCGCAGTGCCTCGCAGCCAGCCAGTTCAGTTCAGGGGCTACCA 866
QY 80 IleCysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAla 99
Db 867 CTGTGGGGGGCTCTGTCATCAGCCCTTGATGATCATCTGCTGCACACCTGTGTTAT 926
QY 100 AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThr 119
Db 927 GACTTGATCCTCCCAAGTCATGAGCATCCAGGTGGGTCTAGTTTCCCTGTGGAC--- 983
QY 120 AspProGlyGluGlnThrIleGluThrValIleIleHisProHisPheSerThr 139
Db 984 AATCCAGCCCCATCCACTTG---GTGGAGAAAGATTGTTTACCACAGCAAGTACAGCCA 1040
QY 140 LysLysProMetAspTyrAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGly 159
Db 1041 AAGAGG---CTGGCAATGACATCGCCCTTATGAGCTGAGTGCAGCCCTCAGTTCAT 1097
QY 160 HisPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIle 179
Db 1098 GAAATGATCCAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
QY 180 CysThrThrAlaGlyTyrPheGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGln 199
Db 1158 TGCTGGAGCTCAGGATGGGGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1217
QY 200 GluValAsnLeuProIleLeuThrTrpGluGlu-CysValAlaAlaLeuLeuThrLeu 219
Db 1218 CACGGCGCGCTCCCTTTGATTTCCAAAGAAAGATCTGC-----AACCA 1259
QY 219 sArgProIleSerGly-----LysThrPheLeuCysThrGlyPheProAspG 235
Db 1260 CAGGACGTGACGTGGCGTCATCTCCCTCCATGCTGCTGCTGCTGCTGCTGCTGCTG 1318
QY 235 yGlyArg---AspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLys 254
Db 1319 GTGGCTTGGAACACGTCCAGGGGGGACAGCGGGGGGGCCCTGGTGTGTCAAGAGAGG 1378
```

RESULT 2

```
US-09-261-416-1
; Sequence 1, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease
; FILE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
```

```

QY 254 sGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpAr 274
Db 1379 G-----CTGTGGAAGTTAGTGGGACGACAGCTTTGGCATGGGCTGC----- 1421
QY 274 gAsnAsnValArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysVa 294
Db 1422 -----GCAGACTGAACAGCTGGGTGGTGTACACCGGTGTACACCTCCCTT 1465
QY 294 lLeuSerTrpIleHisGluHisIleGln 303
Db 1466 CTGGACTGGATCCAGCAGCAGATGGAG 1493

RESULT 3
US-09-386-642-7
; Sequence 7, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: QI, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-7

Alignment Scores:
Pred. No.: 9,63e-41 Length: 1169
Score: 444.50 Matches: 100
Percent Similarity: 49.03% Conservative: 52
Best Local Similarity: 32.26% Mismatches: 119
Query Match: 27.17% Indels: 39
DB: 4 Gaps: 8

US-09-735-713A-2 (1-306) x US-09-386-642-7 (1-1169)

QY 3 LeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhePheGlu 22
Db 46 CTGCTCTGCTGTGTGGTGTCAAAATCTACTCTTGTGCCAGGGTGTGGTCTCCGACTAC 105
QY 23 ArgGly-----LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGly 38
Db 106 AAGGACGACGACGACGTGGGCGGCGCTCTTGTGCGCCCTC----- 147
QY 39 GlnSerLeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGly 58
Db 148 -----TTTGATGATGATGACACAGATCGTTGG 174
QY 59 GlySerGlnValGlnLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLys 78
Db 175 GGATATGCTCTAGAGCGCGGTGAGTGGCCCTGGCAGGTGACATCCTATGAAGCGTC 234
QY 79 HisIleCysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIle 98
Db 235 CATGTGTGGTGGTCTCTCTGCTGTGACAGTGGGTGTGCTGAGTCTCACTCTTC 294
QY 99 AlaAsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGln 118
Db 295 CCCAGCGAGCACCACCAAGGAAGCTATGAGGTCAAGCTGGGGGCCACCACTAGACTCC 354
QY 119 ThrAspProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSer 138
Db 355 TACTCCGAGGACGCCAAGGTGACACCCCTGAGGAGCATCATCCGCCACCCAGCTAC--- 411

```

```

QY 139 ThrLysLysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPhe 158
Db 412 CTCAGGAGGGCTCCCGAGGCGGACATTGCACTCTCCAACTCAGCAGCCCATCACCTTC 471
QY 159 GlyHisPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPhe 178
Db 472 TCCCGCTACATCGGGCGCCATCTCCCTCGCAGCCCAACGCCCTCTTCCCAACGGCTC 531
QY 179 IleCysThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeu-----SerGln 196
Db 532 CACTGCACCTCTCACTGGCTGGGTGATGTGCCCCCTCAGTGAGCCTCTCTGACGCCCAAG 591
QY 197 ValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeu 216
Db 592 CCACTGCAGCAACTCGAGGTGCCTCTGATCAGTCGTGAGACGCTGTAACCTGCCTGACAAC 651
QY 217 ThrLeuLysArgPro-----IleSerGlyLysThrPheLeuCysThrGlyPhe 232
Db 652 ATCGAGGCCCAAGCTCGAGGAGCGCACTTTGTCCAGAGGACATGTTGTGTGTGGCTAT 711
QY 233 ProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsn 252
Db 712 GTGGAGGGGGCAAGGACGCTGCCAGGTGACTCTGGGGGGCCCACTCTCTCTGC---CCT 768
QY 253 LysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGly 272
Db 769 GTGGAGGGTCTCTGGTACCTGACGGGCACTGTGAGCTGGGAGATGCTGTGGGGCC--- 825
QY 273 TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSer 292
Db 826 ---CGCAACAGG-----CCTGGTGTGTACTACTCTGGCCCTC 858
QY 293 LysValLeuSerTrpIleHisGluHisIle 302
Db 859 AGCTATGCTCTCTGGATCCAAAGCAAGGTG 888

RESULT 4
US-09-510-738A-188
; Sequence 188, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 188
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length cDNA of hepsin
US-09-510-738A-188

Alignment Scores:
Pred. No.: 1.83e-40 Length: 1783
Score: 444.50 Matches: 103
Percent Similarity: 49.51% Conservative: 50
Best Local Similarity: 33.33% Mismatches: 120
Query Match: 27.17% Indels: 36
DB: 4 Gaps: 9

US-09-735-713A-2 (1-306) x US-09-510-738A-188 (1-1783)

QY 16 LeuGlyIleValPhePheGluArgGlyLysSerAla----- 27
Db 587 GTGGGCTTCTTCTGTGTGGAGGAGGAGGCTGCCACACCCAGAGGCTGTGGAGGT 646
QY 28 -----AlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGln 45

```



```
US-09-735-713A-2 (1-306) x PCT-US94-00616-1 (1-2581)
QY 36 SerCysGlyGlnSerLeuVal-----LysValGlnProTrpAsnTyrPheAsnIlePhe 53
Db 1649 TCATGTGGGAAACACTGGTACTCAAGAAATTAGCCG-----1687
QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 1688 ---AGATGTCCGAGAAAGTACTCCAGAGAAGGAGCCCTGGGCTGGTGTCTG 1744
QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 1745 TATTTCGACGATCAACAGGTCTCCGAGGCTTCTCTGGTGAGCAGGATGGCTGGTGTG 1804
QY 94 AlaAlaHisCysIleAlaAsnArgAsnIle---ValSerThrLeuAsnValThrAlaGly 112
Db 1805 GCCGCCCACTGCTGACGGAGAAATATGAGCCGCTCTAAGTGGAAAGCAGGTAGGC 1864
QY 113 GluTyrAspLeuSerGln---ThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 1865 CTGCATATGGCATCAATCTGACTTCTCTCAGATAGAACCTAGGTGATTGACCAAT 1924
QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys 151
Db 1925 GTCAATAACCCACACTACAAT---AAACGAGAAAGAACAAATGACATTCGATGCGAT 1981
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 1982 CTGGAATGAAAGTGAACATACACAGATTATATACAGCTATTGTTTACCAGAGAAAT 2041
QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 2042 CAAATTTTCCCCAGGAGAAATTTGTTCTATTGCTGGCTGGGGGCACTTATATATCAA 2101
QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGlyCys 211
Db 2102 GGTCTTACTGCAGACGTACTGCAAGAGCTGAGTTCCTTCCCTCTATCAATGAGAAAT 2161
QY 212 ValAlaAlaLeu-----LeuThrLeuLysArgProIleSerGlyLysThrPhe 227
Db 2162 CAACAACAGATGCCAGAAATATAACATAGC-----GAAAAATATG 2200
QY 228 LeuCysThrGlyPheProAspGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247
Db 2201 GTGTGTGAGGCTATGAGCAGAGGAGGTAGATTCTGTGAGGGGGATTCAGCGCGACCA 2260
QY 248 LeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeu 267
Db 2261 CTCATGTGCCAA---GAAAAACAACAGATGGCTCTGCTGGCTGCGTCACTTTGGATAT 2317
QY 268 GlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIle 287
Db 2318 CAATGTGCACTGCCCTAAATCGC-----CCAGGGGTG 2347
QY 288 PheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302
Db 2348 TATGCCCCGGTCCCAGGTTTCACAGATGGATACAAAGTTTCTA 2392
RESULT 8
US-09-000-846-1
: Sequence 1, Application US/09000846
: Patent No. 5981830
: GENERAL INFORMATION:
: APPLICANT: WU, QINGYU
: APPLICANT: SADLER, JASPER
: TITLE OF INVENTION: KNOCKOUT MICE AND THEIR PROGENY WITH
: NUMBER OF INVENTION: DISRUPTED SERINE PROTEASE GENES
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MILLER, WHITE, ZELANO & BRANTIGAN, P. C.
: STREET: 2200 CLARENDON BLVD, SUITE 1400
: CITY: ARLINGTON
```

```
STATE: VA
COUNTRY: US
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/000,846
FILING DATE: 30-DEC-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/866,058
FILING DATE: 30-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: LEBOVITZ, RICHARD M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: BERLX 65P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: 35..1282
US-09-000-846-1
Alignment Scores:
Pred. No.: 4,42e-40 Length: 1605
Score: 440.50 Matches: 102
Percent Similarity: 49.48% Conservative: 41
Best Local Similarity: 35.29% Mismatches: 102
Query Match: 26.93% Indels: 45
DB: 2 Gaps: 10
US-09-735-713A-2 (1-306) x US-09-000-846-1 (1-1605)
QY 31 LeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPhe 50
Db 471 CTGCCACCTGCCAA-GACTGTGGCGCAGAGAGCTGCCGGTG-----511
QY 51 AsnIlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGln 70
Db 512 -----GACCGCATTTGGGGGCCAGGACAGCACTCTGGGAGGTGGCCGTGGCAG 562
QY 71 ValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrp 90
Db 563 GTCAGCCTGGGTTATGATGGGACCCACCTCTGTGGGGGTCCCTGCTGTCTGGGACTGG 622
QY 91 VallIleThrAlaAlaHisCysIleAlaAsnArgAsn---IleValSerThrLeuAsnVal 109
Db 623 GTGTGACTGTGTCACATTTGCTTTCCAGAGGCGGAACCGGGTCTCTCTCGGTGGCGAGTA 682
QY 110 ThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGlu 129
Db 683 TTTGCTGTGTGCT-----GTAGCCCGGACCTCACCCATGCTGTCAACTGGGGGTTT 736
QY 130 ThrValIleIleHisProHisPhe-----SerThrLysLysProMetAsp 144
Db 737 GCTGTGATCTATCATGGGGGCTACCTTCCCTTTCGAGACCCCTACTATCGACGAAACAGC 796
QY 145 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 164
```

D	b	106	::: GCTCTGTGGCCGC-----	120
Q	y	48	AsnTyrPheAlaPheSerArgIleLeuGlySerGlnValGluLysGlySerTyr ::: ::::: ::: ::: ::: ::	67
D	b	121	-----TTTGATGATGACACAAGATCGTTGGGGCTCATGCTCTAGAGGCCGGTCAGTGG ::: ::::: ::: ::: ::: :::	174
Q	y	68	ProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSer ::: ::::: ::: ::: ::: :::	87
D	b	175	CCCTGGCAGGTCAGCATCACCTATGAAGGGCTCCATGTGTGTGGTCTCTCTCGTGCT ::: ::::: ::: ::: ::: :::	234
Q	y	88	ProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu ::: ::::: ::: ::: ::: :::	107
D	b	235	GAGCAGTGGGTGTGTGAGCTGCTCACTGCTCCCCAGCGAGCACCAAGAAGGCCTAT ::: ::::: ::: ::: ::: :::	294
Q	y	108	AsnValThrAlaGluGluTyAspLysSerGlnThrAspProGlyGluGlnThrLeuThr :::: :::: ::::	127
D	b	295	GAGGTCAAGCTGGGGGCCACCAGCTAGACTCTACTCCGAGNAGCCCAAGGTCAGCAC :::: :::: ::::	354
Q	y	128	IleGluThrValIleIleHisProHisPheSerThrLysLysPrometAspTyrAspIle ::::: :::: :::: ::::	147
D	b	355	CTGAAGGACATCATCCCCACCCCAGGTAC---CTCCAGGAGGGCTCCCAAGGGCGACATT ::::: :::: :::: ::::	411
Q	y	148	AlaLeuLeuLysMetAlaGlyAlaPheGlnPheClyHisPheValGlyProIleCysLeu ::: :::: :::: ::::	167
D	b	412	GCATCTCTCCAACTCAGCAGACCCATCACCTTCTCCCGCTACATCCGGGCCCATCTGCCTC ::: :::: :::: ::::	471
Q	y	168	ProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArg :::: :::: ::::	187
D	b	472	CCTGAGCAACAGCGCTCTCTCCCCAAGCGCTCCACTGCATCTGCTGCTGGGGTCAT :::: :::: ::::	531
Q	y	188	LeuThrGluGlyGlyValLeu-----SerGlnValLeuGlnValAsnLeuProIle :::: :::: ::::	205
D	b	532	GTGGCCCCCTCASTGAGCTCTCTGACGCCCAAGCCATCTSCAGCAACTCGAGGTGCCTCTG ::::: :::: :::: ::::	591

RESULT 9

Qy		LeuIrrpGlnGluCysValArgAlaGluMetThrLeuLysAspPro	220
Db	592	ATCATGTCTGAGACGTGTAACTGCCTGTACAACATCACGCCCAAGCCCTGAGAGCCGCCAC	651
Qy		LeuCysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGln	241
Db	652	TTTGTCGAAGAGACATGGTGTGTGCTGTATGTGAGGGGGCGCAAGAGACGCTGCCAG	711
Qy	242	GlyaspSerGlyGlySerLeuMetCysArgasnLysLysGlyAlaTrpThrLeuAlaGly	261
Db	712	GGTACTCTGGGGGCCCATCTCTCTGC---CCTGTGAGGGTCTCTGGTACCCTGACGGCG	768
Qy	262	ValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgasnValArgLysSerAsp	281
Db	769	ATTGTGACCTGGGAGATGCCTGTGGGCC-----CGCAACAGG-----	807
Qy	282	GlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHis	301
Db	808	-----CCTGGTGTGTACACTCTGCCTCCAGCTATGCCTCCTGGATCCAAGCAAG	858
Qy	302	Ile 302	

Alignment Scores:

RESULT 10
US-09-518-046-3
; Sequence 3, Application US/09518046
; Patent No. 6294663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.

US-09-735-713A-2 (1-306) x US-09-386-642-8 (1-1142)

```

; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof
;
; FILE REFERENCE: D6192CIP
;
; CURRENT APPLICATION NUMBER: US/09/518,046
;
; CURRENT FILING DATE: 2000-03-02
;
; EARLIER APPLICATION NUMBER: 09/261,416
;
; EARLIER FILING DATE: 1999-03-03

```



```
Query Match: 26.56% Indels: 38
DB: 4 Gaps: 13

US-09-735-713A-2 (1-306) x US-09-004-731-32 (1-1200)

QY 9 ArgAsnLysLeuLeuLeuLeuGlyIleValPheGluArgGlyLysSerAlaA 28
DB 366 CGGAATGATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 410
QY 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48
DB 411 GCGCAGCACCA-----GGCGCTCTGCAGAACGC---AACTTGTGGAGA 452
QY 48 nTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyPr 68
DB 453 ATTGTACACAGGCTCTAATAGATCTAGAGGTCTATTCACAGGATTCGGGTCTCATCC 512
QY 68 oTrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyGlySe 84
DB 513 TTGGCAGGGGCTTTGATCAATCAGGATTTTGTAGTAAAAAATTTATCTTGGGTGGCGC 572
QY 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104
DB 573 TTTAGTTAGCATCGATGGGTATATACTGCTGCACATTCGTTGCCACG---ACACCAA 629
QY 104 lSerThrLeuAsnValThrAlaGlyGluTyraSPLeuSerGlnThrAspPro-----G1 122
DB 630 TTCGAACCTGAAGTCCGATTGGCGCAATGGGACGTCCGGACCATGATGAGCGACTGAA 689
QY 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPr 142
DB 690 CCACGAGGAATACGCAATCGAACGCAAGAGTTTCATCTTCATATTCATCA-----CC 740
QY 142 oMetAspTyP-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHi 160
DB 741 AACCGATTTCGGAATGATGATGACCTTAGTGAACCTCGATAGAACCTGTTATTTCAAACA 800
QY 160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy 180
DB 801 ACATATTTACCTGCTGCTTACCTCAT---AAGCAATGAACACTGCTGGAAAAATGGC 857
QY 180 sThrThrAlaGlyTrpGlyArgLeuThrGluGly---GlyValLeuSerGlnValLeuGl 199
DB 858 AACAGTCGCGGATGGGACGAGCAGGACGAGGACGACTGTCCCGGCTGCTTACA 917
QY 199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu-- 218
DB 918 AGAAGTCGATGTGCGAGGTGATTCGGAATGAAGATGCCAGAGGTGTTCCGCTGCTCGGG 977
QY 219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs 238
DB 978 TCACAGAGAACCATTCACGATGCTCTTCTCGCGCGGATATAAAGAGGGGTGCTGTA 1037
QY 238 pAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpTh 258
DB 1038 TTCTATCCCAAGGTGATCTCGAGGTCTCTAATAATGACAG---ATTGAGGTAGAGGAC 1094
QY 258 rLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValAr 278
DB 1095 CCTTGTGGGTCTAGTTTCTTGGGGCATTCGATGTGCTGCT-----1134
QY 278 gLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIl 298
DB 1135 -----GAGCATTTACAGGCGGTATATACCAATATACAAAATTCATACCGGTGAT 1184
QY 298 e 298
DB 1185 C 1185
```

RESULT 12

US-09-004-731-34/c

; Sequence 34, Application US/09004731

; Patent No. 6177258

```
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1200 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-004-731-34

Alignment Scores:
Pred. No.: 1,35e-39 Length: 1200
Score: 434.50 Matches: 110
Percent Similarity: 53.16% Conservative: 50
Best Local Similarity: 36.54% Mismatches: 103
Query Match: 26.56% Indels: 38
DB: 4 Gaps: 13

US-09-735-713A-2 (1-306) x US-09-004-731-34 (1-1200)

QY 9 ArgAsnLysLeuLeuLeuLeuGlyIleValPheGluArgGlyLysSerAlaA 28
DB 835 CGGAATGATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
QY 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48
DB 790 GCGCAGCACCA-----GGCGCTCTGCAGAACGC---AACTTGTGGAGA 749
QY 48 nTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyPr 68
DB 748 ATTGTACACAGGCTCTAATAGATCTAGAGGTCTATTCACAGGATTCGGGTCTCATCC 689
QY 68 oTrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyGlySe 84
DB 688 TTGGCAGGGGCTTTGATCAATCAGGATTTTGTAGTAAAAAATTTATCTTGGGTGGCGC 629
QY 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104
DB 628 TTTAGTTAGCATCGATGGGTATATACTGCTGCACATTCGTTGCCACG---ACACCAA 572
QY 104 lSerThrLeuAsnValThrAlaGlyGluTyraSPLeuSerGlnThrAspPro-----G1 122
```


29	LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValGlnProTrp--As	48
Db	::: :::: :::: :::: :::: ::::	
790	GCSCGACACCA-----GGSCGCTCTGCAGAACGC---AACTTGTGCGAGA	749
Qy	48 nTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrPr	68
Db	::: :::: :::: :::: :::: :::: ::::	
748	ATTGTACACGAGGTCTAATAAGAAATCGTAGAGGTCAATCAACAGGATTTCGGGTCTCATCC	689
Qy	68 oTTPGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyGlySe	84
Db	:::: :::: :::: :::: :::: ::::	
688	TTGCGACCGCGCTTGATCAATCAAGATTTTGTAGTAAAAAATATCTTTCGGGTGGCGC	629
Qy	84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal	104
Db	::: :::: :::: :::: :::: :::: ::::	
628	TTTAGTTAGCGATCGATGGGTATTAAGTCTGCACATTCGGTTGCCACG---ACACCAA	572
Qy	104 lSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspPro-----G	122
Db	::: :::: :::: :::: :::: :::: ::::	
571	TTTCGAACCTGAAGTGCATTTGGCGAATGGGACGTCGCGACACCGATCAGGCGACTGAA	512
Qy	122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPr	142
Db	::: :::: :::: :::: :::: :::: ::::	
511	CCACGAGGAATACGCAATCAACGCAAGAAAGTTCATCTCTCATATTCA-----CC	461
Qy	142 oMetAspTyr-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyH	160
Db	:::: :::: :::: :::: :::: ::::	
460	AACGGATTTCCGGAATGATGACCTTAGTGAACACTCGATAGAACCTGTATTATTTCAACA	401
Qy	160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy	180
Db	::: :::: :::: :::: :::: :::: ::::	
400	ACATATTTTACCTGCTGCTTACCTCAT---AAGCAAAATCAAACTGCCTGGAATAATGCC	344
Qy	180 sThrThrAlaGlyTrpGlyArgLeuThrGluGly-----GlyValLeuSerGlnValLeuG	199
Db	:::: :::: :::: :::: :::: ::::	
343	AACAGTCCCGCGATGGGACGCGACGAGGCACGGCGAGACGACTGTCCGCGCTCTCTTACA	284
Qy	199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu--	218
Db	:::: :::: :::: :::: :::: ::::	
283	AGAAAGTCGATGTCGAGGTGATCCGGAATGAAGATGCCAGAGGTGGTTCGCTGCTCGGG	224
Qy	219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs	238
Db	::: :::: :::: :::: :::: :::: ::::	
223	TCGACGAGAAACCATTCACCATGCTCTTCTCTCGCGCGGATATAAGAGGGTGGTCTGTA	164
Qy	238 pAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaTrpTh	258
Db	:::: :::: :::: :::: :::: ::::	
163	TTATGTCGAAGGTGATTCTGGAGGTCTCTTAATAATGCAG---ATTGAGGGTGAAGAGGAC	107
Qy	258 rLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValar	278
Db	:::: :::: :::: :::: :::: ::::	
106	CCCTGTGGGTCTAGTTCTTTGGGCGATGTGATGTGTGCTGCT-----	67
Qy	278 gLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIl	298
Db	::: :::: :::: :::: :::: :::: ::::	
66	-----GAGCAATTACAGCGGTATATACCAATATACAAAAATTCATACCGTGGAT	17
Qy	298 e 298	
Db	16 C 16	

```

RESULT 15
US-09-004-729-32
; Sequence 32, Application US/09004729
; Patent No. 6406300
; GENERAL INFORMATION:
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Stiedler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEAS
; TITLE OF INVENTION: MOLECULES AN
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:

```

Qy 9 ArgAsnLysLeuIleLeuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAla 28
||||| |||:::|||| ::||| |||||
Db 835 CGAATGATCTGGTGGCTCGCTCGGT-----CGACAGGAGCATTTCGGCCTGA 79

ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,729
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1200 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1200
US-09-004-729-32

Alignment Scores:
Pred. No.: 1.35e-39 Length: 1200
Score: 434.50 Matches: 110
Percent Similarity: 53.16% Conservative: 50
Best Local Similarity: 36.54% Mismatches: 103
Query Match: 26.56% Indels: 38
DB: 4 Gaps: 13

US-09-735-713A-2 (1-306) x US-09-004-729-32 (1-1200)

Qy 9 ArgAsnLysLeuIleLeuLeuGlyIleValPhePheGlyLysSerAlaAla 28
Db 366 CGAATGATCTGGTCTGCTGCGT-----CGACAGGAGCATTCGGCCTGA 410
Qy 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48
Db 411 GCCGACGACCA-----GGCCCTCTGCAGAACGC---AACTTGTGGAGA 452
Qy 48 nTyrPheAsnIlePheSerArgIleLeuGlySerGlnValGluLysGlySerTyPr 68
Db 453 ATTGTACAGAGGTCTAATAGAATCGTAGAGTTCATTCAACAGGATTCGGGTCTCATCC 512
Qy 68 cPrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyLys 84
Db 513 TTGGCAGCGGCTTTGATCAATCATAGGATTTTGTAGTAAAAAATTATCTGCGGTGGCG 572
Qy 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104
Db 573 TTTAGTAGCATGATCGGTTATACCTGCTGCACATTCGCTGCCACG---ACACAAA 629
Qy 104 lSerThrLeuAsnValThrAlaGlyGluTyPrAspLeuSerGlnThrAspPro-----G1 122
Db 630 TTCGAACCTGAAGGTGCGATTGGCGAATGGGACGTCGCGACCCACGATGACGACTGAA 689
Qy 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisProHisThrLysLysPr 142
Db 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisProHisThrLysLysPr 142

..

Db 690 CCACGAGGAATACGCAATCGAAACGCAAGAGTTTCATCTTCATATTCA-----CC 740
Qy 142 oMetAspTyr-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHi 160
Db 741 AACCGATTTCCGAATGATGACCTTAGTGAACCTCGATAGAACTGTTATTTTCAACA 800
Qy 160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy 180
Db 801 ACATATTTTACCTGCTGCTTACCTCAT---AAGCAATGAAACTGGCTGGAATAATGCG 857
Qy 180 sThrThrAlaGlyTrpGlyArgLeuThrGluGly---GlyValLeuSerGlnValLeuGl 199
Db 858 AACAGTCGCGGATGGGACGAGCAGGACGAGCAGCTGTGCCGGCTGCTTACA 917
Qy 199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu-- 218
Db 918 AGAAGTCGATGTCGAGGTGATTCGGAATGAAAGATGCCAGAGTGGTTCGCTGCGGG 977
Qy 219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs 238
Db 978 TCGACGAGAAACCAATTCACGATGCTCTCTGCGCGGATATAAAGAGGTGGTGTGCTGA 1037
Qy 238 pAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpTh 258
Db 1038 TTCATGCCAAGGTGATTCGAGGTCTCTAATAATGCAG---ATTGAGGCTAGAGAGGAC 1094
Qy 258 rLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValAr 278
Db 1095 CTTTGGGTCTAGTTCTTGGGGCATTTGGATGGTCTG-----1134
Qy 278 gLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpI 298
Db 1135 -----GAGCATTTTACCAGGCGTATATACCAATATACAAAAATTCATACCGTGGAT 1184
Qy 298 e 298
Db 1185 C 1185

Search completed: March 25, 2003, 02:04:14
Job time : 66 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:48:49 ; Search time 20 Seconds
(without alignments)

634.588 Million cell updates/sec

Title: US-09-735-713A-2

Perfect score: 1636

Sequence: 1 MSUKMLISRNKLLLLGIVF.....IFTDISKVLSWIHEHIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	486	29.7	786	1	STUB_DROME
2	466	28.5	638	1	KAL_HUMAN
3	464	28.4	638	1	KAL_MOUSE
4	456.5	27.9	454	1	TMS3_HUMAN
5	447.5	27.4	638	1	KAL_RAT
6	446	27.3	342	1	PSS8_MOUSE
7	446	27.3	342	1	PSS8_RAT
8	444	27.1	343	1	PSS8_HUMAN
9	442.5	27.0	417	1	HEPS_HUMAN
10	442.5	27.0	1035	1	ENTK_BOVIN
11	440.5	26.9	416	1	HEPS_MOUSE
12	438.5	26.8	416	1	HEPS_RAT
13	438.5	26.8	1034	1	ENTK_PIG
14	438	26.8	457	1	TMS5_HUMAN
15	433	26.5	593	1	FAL2_BOVIN
16	431.5	26.4	1019	1	ENTK_HUMAN
17	429	26.2	290	1	MPN_HUMAN
18	427	26.1	311	1	TRYG_MOUSE
19	426	26.0	855	1	ST14_MOUSE
20	425	26.0	1069	1	ENTK_MOUSE
21	423	25.9	455	1	TMS5_MOUSE
22	421.5	25.8	317	1	BSS4_HUMAN
23	421.5	25.8	321	1	TRYG_HUMAN
24	419.5	25.6	237	1	TRYP_ASTFL
25	419	25.6	264	1	CTRL_HUMAN
26	419	25.6	273	1	TRYT_SHEEP
27	418	25.6	306	1	BSS4_MOUSE
28	418	25.6	625	1	FALL1_HUMAN
29	415.5	25.4	275	1	TRY3_ANOGA
30	414.5	25.3	615	1	FAL2_HUMAN
31	413.5	25.3	263	1	CTRL_CANFA
32	413.5	25.3	274	1	TRY1_ANOGA
33	413	25.2	437	1	TMS4_HUMAN

ALIGNMENTS

RESULT 1

ID	STUB_DROME	STANDARD;	PRT;	786 AA.
AC	Q05319;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Serine proteinase stubble (EC 3.4.21.-) (Stubble-stubbloid protein).			
GN	SB OR SB-SBD			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=727;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Oregon-R;			
RX	MEDLINE=93281671; PubMed=7685111;			
RA	Appel L.F., Prout M., Abu-Shumays R., Hammonds A., Garbe J.C.,			
RA	Fristrom D., Fristrom J.;			
RT	"The Drosophila Stubble-stubbloid gene encodes an apparent			
RT	transmembrane serine protease required for epithelial			
RT	morphogenesis."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:4937-4941(1993).			
CC	-!- FUNCTION: HORMONE DEPENDENT PROTEASE REQUIRED FOR EPITHELIAL			
CC	MORPHOGENESIS. HAS A DUAL FUNCTION, DETACHES IMAGINAL DISC CELLS			
CC	FROM EXTRACELLULAR MATRICES THROUGH ITS EXTRACELLULAR PROTEOLYTIC			
CC	DOMAIN AND TRANSMITS AN OUTSIDE-TO-INSIDE SIGNAL TO ITS			
CC	INTRACELLULAR DOMAIN TO MODIFY THE CYTOSKELETON DURING			
CC	MORPHOGENESIS. MAY BE ABLE TO ACTIVATE ITSELF.			
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.			
CC	-!- INDUCTION: BY 20-HYDROXYECYDSONE (20HE).			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-24 IS THE INITIATOR.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L11451; AAA28918.1; -			
DR	HSPSP; P00763; 1DPO.			
DR	MEFOPS; S01.225; -			
DR	FlyBase; FBgn0003319; Sb.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin.1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_Spc; 1.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Transmembrane; Glycoprotein; Zymogen;			

34	409	25.0	263	1	CTRB_HUMAN	PI7538 homo sapien
35	409	25.0	275	1	TRY4_ANOGA	P35038 anopheles g
36	408.5	25.0	415	1	ACRO_PIG	P08001 sus scrofa
37	408.5	25.0	603	1	FAL2_CAVPO	Q04962 cavia porce
38	408	24.9	490	1	TMS2_MOUSE	Q91iq8 mus musculu
39	407.5	24.9	459	1	PRTC_PIG	Q9qlp2 sus scrofa
40	407	24.9	324	1	TEST_MOUSE	Q9jh17 mus musculu
41	407	24.9	492	1	TMS2_HUMAN	O15393 homo sapien
42	406	24.8	234	1	TRY3_AEDAE	P29786 aedes aegyp
43	404	24.7	407	1	FAT7_BOVIN	P22457 bos taurus
44	403	24.6	314	1	TEST_HUMAN	Q9y6m0 homo sapien
45	402.5	24.6	275	1	TRYT_PIG	Q9n2d1 sus scrofa

```

KW Signal-anchor.
FT CHAIN 1 542 NON-CATALYTIC CHAIN (POTENTIAL).
FT CHAIN 543 786 CATALYTIC CHAIN (POTENTIAL).
FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 59 80 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 81 786 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 543 786 SERINE PROTEASE.
FT DOMAIN 267 276 POLY-SER.
FT DOMAIN 287 298 POLY-GLN.
FT DOMAIN 391 478 SER/THR-RICH.
FT DOMAIN 412 422 POLY-THR.
FT DOMAIN 471 478 POLY-SER.
FT ACT_SITE 589 589 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 639 639 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 737 737 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 531 659 INTERCHAIN (BY SIMILARITY).
FT DISULFID 574 590 BY SIMILARITY.
FT DISULFID 703 722 BY SIMILARITY.
FT DISULFID 733 762 BY SIMILARITY.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 786 AA; 85010 MW; CE3E755760B9DE4D CRC64;

Query Match 29.7%; Score 486; DB 1; Length 786;
Best Local Similarity 37.5%; Pred. No. 1.6e-37;
Matches 98; Conservative 50; Mismatches 87; Indels 26; Gaps 8;

QY 54 SRTLGSGVQKGVQVSLKQ-----RQKHICGSGIVSPQWVITAACHIANRNIVSTL 107
DB 541 TRIVGKSAFGRWPHQVSVRTSPFGSSFTRCGGALINENWATAGHCVDLD-LISQI 599
QY 108 NVTAGYDLSQDTPGQDTIE-----TVIHPHSTKPKMDYDIALLKMGAFQFGHFY 163
DB 600 RIRVGEDFSHQV--BQLPYIERGVAKVHVHPKYSF-LTYEYDLALVKLEPLFAPHYS 656
QY 164 PICLPRLRQFAGFCTTAGRGLTEGVLSQVLOEVNLPILTWECEVAALLTL-KRPI 222
DB 657 PICLPE-TDLSLLGMNATVWGRSEGGTLPVLSQEVSVPIVSDNCKSMFMFRAGRQBF 715
QY 223 SGKTFELTFPPDGGDACQDGGSLMCRNKKGAWTLAGVTSWGLCGGRGNVNRKSDQ 282
DB 716 IPDIFLCAGYETGGQDSCQDGGSLPQAKSQDGRFLAGIISWGIGC-----AEA 765
QY 283 GSPGIFTDLSKVLNWIHQ 303
DB 766 NLPGVCTRISKETPWILEHR 786

RESULT 2
KAL_HUMAN STANDARD; PRT; 638 AA.
AC P03952;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (Kininogenin) (Fletcher factor).
GN KLKB1 OR KLK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86243359; PubMed=3521732;
RA Chung D.W., Fujikawa K., McMullen B.A., Davie E.W.;
RT "Human plasma prekallikrein, a zymogen to a serine protease that
RT contains four tandem repeats.";
RL Biochemistry 25:2410-2417(1986).
RN [2]
RP SEQUENCE FROM N.A.; AND VARIANTS SER-143; GLN-202 AND PRO-208.
RX MEDLINE=20487549; PubMed=11031105;
RA Yu H., Anderson P.J., Freedman B.I., Rich S.S., Bowden D.W.;

```

```

RT "Genomic structure of the human plasma prekallikrein gene,
RT identification of allelic variants, and analysis in end-stage renal
RT disease.";
RL Genomics 69:225-234(2000).
RN [3]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=91152016; PubMed=1998666;
RA McMullen B.A., Fujikawa K., Davie E.W.;
RT "Location of the disulfide bonds in human plasma prekallikrein: the
RT presence of four novel apple domains in the amino-terminal portion of
RT the molecule.";
RL Biochemistry 30:2050-2056(1991).
CC -!- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-I-Xaa and Lys-I-Xaa
CC bonds, including Lys-I-Arg and Arg-I-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -!- DISEASE: Defects in KLKB1 are the cause of Fletcher factor
CC deficiency, a blood coagulation defect.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 4 APPLE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13143; AAA60153.1; .
DR EMBL; AF232742; AAF79940.1; .
DR EMBL; AF232734; AAF79940.1; JOINED.
DR EMBL; AF232735; AAF79940.1; JOINED.
DR EMBL; AF232736; AAF79940.1; JOINED.
DR EMBL; AF232737; AAF79940.1; JOINED.
DR EMBL; AF232738; AAF79940.1; JOINED.
DR EMBL; AF232739; AAF79940.1; JOINED.
DR EMBL; AF232740; AAF79940.1; JOINED.
DR EMBL; AF232741; AAF79940.1; JOINED.
DR PIR; A00921; KQHUP.
DR PIR; A37939; A37939.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.212; .
DR Genew; HGNC:6371; KLKB1.
DR MIM; 229000; .
DR InterPro; IPR000177; Apple.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR000005; APPLEDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00223; APPLE; 4.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00495; APPLE; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat; Polymorphism.
FT SIGNAL 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.

```

```
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .).
FT ACT_SITE 434 434 ACT_SITE
FT ACT_SITE 483 483 ACT_SITE
FT ACT_SITE 578 578 ACT_SITE
FT DISULFID 21 104 CHARGE RELAY SYSTEM.
FT DISULFID 47 77 CHARGE RELAY SYSTEM.
FT DISULFID 51 57
FT DISULFID 111 194
FT DISULFID 137 166
FT DISULFID 141 147
FT DISULFID 201 284
FT DISULFID 227 256
FT DISULFID 231 237
FT DISULFID 292 375
FT DISULFID 318 347
FT DISULFID 322 328
FT DISULFID 340 345
FT DISULFID 383 503
FT DISULFID 419 435
FT DISULFID 517 584
FT DISULFID 548 563
FT DISULFID 574 602
FT VARIANT 143 143
FT VARIANT 202 202 N -> S (COMMON POLYMORPHISM).
FT VARIANT 208 208 /FTid-VAR_013598.
FT VARIANT 208 208 H -> Q.
FT VARIANT 208 208 /FTid-VAR_013599.
FT VARIANT 208 208 H -> P.
FT VARIANT 208 208 /FTid-VAR_013600.
FT SEQUENCE 638 AA; 71369 MW; E62F9C1053838FB4 CRC64;

Query Match 28.5%; Score 466; DB 1; Length 638;
Best Local Similarity 35.5%; Pred. No. 9.4e-36;
Matches 91; Conservative 51; Mismatches 96; Indels 18; Gaps 5;

QY 54 SRILGGQVEKSGVPMQVSLKQ---QKHLGGSGIVSPQVITAAHCIANRNIVSTLNVT 110
DB 389 TRIVGGTNSWGEMPMQVSLQVLTAAQRHLGGSLGHQWLVTAACHFDGLPLQDVWRIY 448
QY 111 AGEVDLSQTDPRGEQTLTIETVTHPHFSTKKPMDDYDIALKMGAFQFHFVGPICLP 170
DB 449 SGILNLSDIITKDPFQSIKEIIHQYKVSEGNHDIALLIKLOAPLNYTEFPICLP 507
QY 171 REQFEAGFICTTAGWRLTEGGVLSQVLEVNLPILTWEECVAAALLFLKRPISGKTF 230
DB 508 GDTSTIYTNWVTGWGFSKEGEIQNLQKVNPLVTNEECQKRYQDYK---ITQRMVCA 564
QY 231 GFDGGRDAGCGSGSLMCRNKGAWTLAGVTSWGLCGCGRWNNVYKSDQSGSPGLFTD 290
DB 565 GYEGGKDGKDGSGGGLPVCKH-NGMRLVGYITSWGEGCAR-----REQPGVYTK 613
QY 291 ISKVLWISHEIOTGN 306
DB 614 VAEMDWILEKTQSSD 629

RESULT 3
ID KAL_MOUSE STANDARD; PRT; 638 AA.
AC P26262;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
```

```
DE (Kininogenin) (Fletcher factor).
GN KLKB1 OR KLK3 OR PK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BA1B/C; TISSUE=Liver;
RX MEDLINE=91090844; PubMed=2264928;
RA Seidah N.G., Sawyer N., Hamelin J., Mion P., Beaubien G.,
RA Brachpapa L., Rochemont J., Mbikay M., Chretien M.;
RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,
RT and comparison of protein and mRNA levels among species.";
RL DNA Cell Biol. 9:737-748(1990)
CC -|- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -|- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -|- SIMILARITY: CONTAINS 4 APPLE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58588; AAA63393.1; -.
DR PIR; A36557; KOMSPL.
DR HSP; P00750; LRTF.
DR MEROPS; S01.212; -.
DR MGD; MGI:102849; Klk3.
DR InterPro; IPR000177; Apple.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00005; APPLEDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00223; APPLE; 4.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00495; APPLE; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat..
FT SIGNAL 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 O-LINKED (PROBABLE).
```



```

DR MM: 605316;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001190; Src_receptor.
DR Pfam: PF00057; ldl_recept_a; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00192; LDLa; 1.
DR SMART: SM00202; SR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS01209; LDLRA_1; 1.
DR PROSITE: PS50068; LDLRA_2; 1.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS50287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Transmembrane; Signal-anchor; Deafness;
KW Alternative splicing; Disease mutation; Polymorphism.
FT DOMAIN 1 48
FT TRANSMEM 49 69
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT LDL-RECEPTOR CLASS A.
FT SRCR.
FT SERINE PROTEASE.
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CLEAVAGE (POTENTIAL).
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT MISSING (IN ISOFORM B/C).
FT EMIOQVCLPNSSEFDPKVCWTSGKATGAGDASPVLN
HAAYPLISNKHNRDVGIIISPMICAGYLGGVDSOG
DSGGPLVQERRUHLVATSGFGICAEVKNKPGVYTRVTSF
LDWIHQEMERDLKT -> GTSGSLCGSAALPLFQEDLQLLI
EAPL (IN ISOFORM D).
FT DLYLPKSWTIQVGLVSLDDNPAPSHLVEKIVYH -> EIVA
PRERADRGKLLCWKPKTKMGKPRPSHS (IN
TRUNCATED ISOFORM).
FT MISSING (IN TRUNCATED ISOFORM).
FT V -> I.
FT /FTid-var_010781.
FT D -> G (IN DFNB8/DFNB10).
FT /FTid-var_013490.
FT R -> W (IN DFNB8/DFNB10).
FT /FTid-var_011678.
FT G -> S.
FT /FTid-var_013491.
FT /FTid-var_013492.
FT D -> N.
FT /FTid-var_013493.
FT C -> F (IN DFNB8/DFNB10).
FT /FTid-var_013494.
FT W -> C (IN DFNB8/DFNB10).
FT /FTid-var_011678.
FT I -> V (IN DFNB8/DFNB10).
FT /FTid-var_013101.
FT P -> L (IN DFNB8/DFNB10).
FT /FTid-var_011679.
FT C -> R (IN DFNB8/DFNB10).
FT /FTid-var_013495.
FT A -> T.

```

```

FT CONFLICT 46 54 /FTid-var_013496.
FT LKFFPIVI -> FEVFSQSSSL (IN REF. 1).
FT A -> T (IN REF. 1).
FT MISSING (IN REF. 1).
FT CONFLICT 350 350 MISSING (IN REF. 1).
FT CONFLICT 369 395 ICHNRDVGIIISPMICAGYLGGVD -> DLQPOGRVRW
FT CONFLICT 427 427 HHLRHALRGLPGWRWN (IN REF. 1).
FT CONFLICT 427 427 E -> D (IN REF. 1).
FT SEQUENCE 454 AA; 49404 MW; 57ECC3678F7D6AFF CRC64;

Query Match 27.9%; Score 456.5; DB 1; Length 454;
Best Local Similarity 32.0%; Pred. No. 4.8e-35;
Matches 99; Conservative 61; Mismatches 114; Indels 35; Gaps 10;

QY 1 MSLKMLISRNKLLILLGIVFERG-KSAALSLPRAPSGOSLVKVPWNYENFISRLGG 59
DB 170 VSDHLLPDDKVTALHHSVYVRECCASGHVVTLOCTAGHRR-----GYSRIVGG 220
QY 60 SQVEKSGYPMQVSLKQKHCIGGSIYSPQWVITAAHCIAHNRNIVSTLVNVTAGYDLSQT 119
DB 221 NMSLSQWPQASLQFQYHLCGGSVITPLWITAAHCVDLYLPKSWTIQVGLVSLDD- 279
QY 120 DPGQQTITETVTHPHSTKPKMDYDIALKMGAFQFHFVGPICLPRLRQEPAGFI 179
DB 280 NPAPSHL-VEKIVYHSKYKPKR-LGNDIALMKLAGPLTFNEMIQPVCLPNSSENFPPDGKV 337
QY 180 CTTAGWGLTEG-GVLSQVLOEVNLPILTWEECVAALLTLKRPISG----KTFCLCTGPPD 234
DB 338 CWTSGWATEDPDGADSPVLNHAAPLISNKKIC-----NHRDVGIIISPMICAGYL 391
QY 235 GGRDACGDSGSLMCRNKGAWTLGAVTSMGLCGGRGNVRNVRKSDQSGPFIETDISKV 294
DB 392 GGVDSGCGSGGLVQCERR-LWKLVGATSEFGICG-----AEVKNKGVYTRVTSF 440
QY 295 LSWTHEHIQ 303
DB 441 LDWTHEQME 449

RESULT 5
KAL_RAT STANDARD; PRT; 638 AA.
ID KAL_RAT
AC P14272;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (Kininogenin) (Fletcher factor).
GN KLKB1 OR PK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91129236; PubMed=1993180;
RX Beaudien G., Rosinski-Chupin I., Mattei M.-G., Mbikay M., Chretien M.,
RA Seidah N.G.;
RT "The cDNA structure of rat plasma kallikrein.";
RL DNA 8:563-574(1989).
CC - FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC - CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)

```

kininogen to release bradykinin.

-! SUBUNIT: THE ZYMAGEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE, AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.

-! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN SUBFAMILY.

-! SIMILARITY: CONTAINS 4 APPLE DOMAINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR	EMBL; M62357; AAA74563.1; -.	1	19	PLASMA KALLIKREIN HEAVY CHAIN.
DR	EMBL; M62358; AAA74563.1; JOINED.	20	390	PLASMA KALLIKREIN LIGHT CHAIN.
DR	EMBL; M62346; AAA74563.1; JOINED.	391	638	APPLE 1.
DR	EMBL; M62347; AAA74563.1; JOINED.	20	105	APPLE 2.
DR	EMBL; M62349; AAA74563.1; JOINED.	110	195	APPLE 3.
DR	EMBL; M62350; AAA74563.1; JOINED.	200	285	APPLE 4.
DR	EMBL; M62351; AAA74563.1; JOINED.	291	376	SERINE PROTEASE.
DR	EMBL; M62352; AAA74563.1; JOINED.	389	621	N-LINKED (GLCNAC. . .) (PROBABLE) .
DR	EMBL; M62353; AAA74563.1; JOINED.	127	127	N-LINKED (GLCNAC. . .) (PROBABLE) .
DR	EMBL; M62354; AAA74563.1; JOINED.	308	308	N-LINKED (GLCNAC. . .) (PROBABLE) .
DR	EMBL; M62355; AAA74563.1; JOINED.	396	396	N-LINKED (GLCNAC. . .) (PROBABLE) .
DR	EMBL; M62356; AAA74563.1; JOINED.	453	453	N-LINKED (GLCNAC. . .) (PROBABLE) .
DR	EMBL; M62357; AAA74563.1; JOINED.	494	494	N-LINKED (GLCNAC. . .) (PROBABLE) .
DR	EMBL; M62358; AAA74563.1; JOINED.	434	434	CHARGE RELAY SYSTEM.
DR	EMBL; M58590; AAA42069.1; -.	483	483	CHARGE RELAY SYSTEM.
DR	PIR; A39180; KQRTPL.	578	578	CHARGE RELAY SYSTEM.
DR	HSSP; P00750; IRTF.	21	104	BY SIMILARITY.
DR	MEROPS; S01.212; -.	47	77	BY SIMILARITY.
DR	InterPro; IPR001177; Apple.	51	57	BY SIMILARITY.
DR	InterPro; IPR001314; Chymotrypsin.	111	194	BY SIMILARITY.
DR	InterPro; IPR003014; PAN.	141	167	BY SIMILARITY.
DR	InterPro; IPR001254; Ser_protease_Try.	147	147	BY SIMILARITY.
DR	Pfam; PF00024; PAN; 4.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00005; APPLIEDOMAIN.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00223; APPLE; 4.			
DR	SMART; SM00020; Tryp_SPC; 1.			
DR	PROSITE; PS00495; APPLE; 4.			
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;			
KW	Fibrinolysis; Blood coagulation; Inflammatory response; Liver;			
KW	Repeat.			
FT	SIGNAL	1	19	
FT	CHAIN	20	390	
FT	CHAIN	391	638	
FT	DOMAIN	20	105	
FT	DOMAIN	110	195	
FT	DOMAIN	200	285	
FT	DOMAIN	291	376	
FT	DOMAIN	389	621	
FT	CARBOHYD	127	127	
FT	CARBOHYD	308	308	
FT	CARBOHYD	396	396	
FT	CARBOHYD	453	453	
FT	CARBOHYD	494	494	
FT	ACT_SITE	434	434	
FT	ACT_SITE	483	483	
FT	ACT_SITE	578	578	
FT	DISULFID	21	104	
FT	DISULFID	47	77	
FT	DISULFID	51	57	
FT	DISULFID	111	194	
FT	DISULFID	137	167	
FT	DISULFID	141	147	

FT	DISULFID	201	284		BY SIMILARITY.
FT	DISULFID	227	256		BY SIMILARITY.
FT	DISULFID	231	237		BY SIMILARITY.
FT	DISULFID	292	375		BY SIMILARITY.
FT	DISULFID	318	327		BY SIMILARITY.
FT	DISULFID	322	328		BY SIMILARITY.
FT	DISULFID	340	345		BY SIMILARITY.
FT	DISULFID	383	503		BY SIMILARITY.
FT	DISULFID	419	435		BY SIMILARITY.
FT	DISULFID	517	584		BY SIMILARITY.
FT	DISULFID	548	563		BY SIMILARITY.
FT	DISULFID	574	602		BY SIMILARITY.
SQ	SEQUENCE	638 AA;	71273 MW;	454BBE27B8CA8F88	CRC64;

Query Match	27.4%;	Score 447.5;	DB 1;	Length 638;	
Best Local Similarity	33.3%;	Pred. No. 5e-34;			
Matches 97;	Conservative	49;	Mismatches 118;	Indels 27;	Gaps 7;

QY	18	IVFFERKSA-ALS	PKAPSCQS	LVKVPW	PNFYFI	SRILGSGVEK	SGYPWQVSL	KQ 76				
DB	360	ITYEAQSGSGY	SLRCKV	SSDCTTKN	-----	ARIVGGTN	SSLGEPWQV	SLQVK 411				
QY	77	---QKH	CGGSIV	PQWVITA	AACHIAN	RNVSTL	NVTAGEY	DL	SOTD	PEQTL	ETVII 133	
DB	412	LVSONHMC	GGSIGT	IGRWILTA	AHCRD	FGIPY	PDVWRIY	SGILNL	SEITNK	TPFSSIKELII 471		
QY	134	HPHSTKK	PMYD	YDITAL	KKAGAF	GFHVG	PICLPEL	REQFEA	GFICT	WAGRL	TEGGV 193	
DB	472	HOKYK	MSG-	SYDIAL	IKLQ	PLNTY	TEPK	PCIP	LSKAD	TNTI	TNCWVT	GMVTKERGE 530
QY	194	LSQVL	ENYLP	LITW	EECA	ALLTL	KRIP	ISCKT	FLCTG	PDGGR	DACQD	SGSGSLMCRNK 253
DB	531	TQNILQ	KATIP	LVN	EECQK	---KYR	DVYIT	KQMIC	AGYKEGG	IDACKD	SGGSLV	PKH- 586
QY	254	KGANT	LACV	TSWGL	CGGR	WNNRV	KKSDQ	SGPI	GTFID	ISKVLS	WIHEIQT 304	
DB	587	SGRMQ	LVGITS	WGE	GCAR	-----	KEQ	PVYTK	VAEYID	WILEKIQS 627		

RESULT 6	
PSS8_MOUSE	STANDARD; PRT; 342 AA.
ID	PSS8_MOUSE
AC	Q9ESDI;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Prostasin precursor (EC 3.4.21.-) (Channel activating protease 1).
GN	PSS8 OR CAP1.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=2035202; PubMed=10770960;
RA	Viugniaux G., Vallet V., Jaeger N.F., Pfister C., Bens M., Farman N.,
RA	Courtois-Coutry N., Vandewalle A., Rossier B.C., Hummler E.;
RT	"Activation of the amiloride-sensitive epithelial sodium channel by
RT	the serine protease mCAP1 expressed in a mouse cortical collecting
RL	duct cell line."
RL	J. Am. Soc. Nephrol. 11:828-834(2000).
CC	-1- FUNCTION: POSSESSES A TRYPsin-LIKE CLEAVAGE SPECIFICITY (BY
CC	SIMILARITY). ACTIVATES AMILORIDE-SENSITIVE SODIUM CHANNELS.
CC	-1- SUBUNIT: HETERODIMER OF TWO CHAINS, LIGHT AND HEAVY, HELD BY A
CC	DISULFIDE BOND (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF
CC	ITS C-TERMINUS (BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC	FRAMESHIFT IN POSITION 339.

CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	at EMBL, Heidelberg.


```

RN CHARACTERIZATION.
RX MEDLINE=91358502; PubMed=1885621;
RA Tsuji A., Torres-Rosado A., Aral T., le Beau M.M., Lemons R.S.,
RA Chou S.H., Kurachi K.;
RT "Hepsin, a cell membrane-associated protease. Characterization,
RT tissue distribution, and gene localization.";
RL J. Biol. Chem. 266:16948-16953(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93348237; PubMed=8346233;
RA Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
RT "Hepsin, a putative cell-surface serine protease, is required for
RT mammalian cell growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).
CC -1- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE
CC OF CELL MORPHOLOGY.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: PRESENT IN MOST TISSUES, WITH THE HIGHEST
CC LEVEL IN LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M18930; AAA36013.1; -.
DR EMBL; X07732; CAA30558.1; -.
DR EMBL; X07002; CAA30058.1; -.
DR PIR; S00845; S00845.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.224; -.
DR Genew; HGNC:5155; HPN.
DR MIM; 142440; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Src_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Transmembrane; Signal-anchor.
FT CHAIN 1 162 NON-CATALYTIC CHAIN (POTENTIAL).
FT CHAIN 163 417 CATALYTIC CHAIN (POTENTIAL).
FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 18 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 45 417 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 163 417 SERINE PROTEASE.
FT ACT_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 257 257 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 353 353 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 153 277 INTERCHAIN (BY SIMILARITY).
FT DISULFID 188 204 BY SIMILARITY.
FT DISULFID 322 338 BY SIMILARITY.
FT DISULFID 349 381 BY SIMILARITY.
FT CARBOHYD 112 112 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 417 AA; 45011 MW; B2086FF661E551D7 CRC64;
Query Match 27.0%; Score 442.5; DB 1; Length 417;
Best Local Similarity 33.0%; Pred. No. 8.8e-34;
Matches 100; Conservative 48; Mismatches 120; Indels 35; Gaps 9;

```

```

QY 11 KLILLGIVFFERCKSAALSAPKAPCSQSLVKVPQWNNFNPISRIILGGSGVSPWQ 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 130 RLLEVISVCDPCPRGRFLAAI---CQDCGRRKLPV-----DRIVGGRTSLGRWPWQ 177

```

```

QY 71 VSLKQROKHICGGSSIVSPQWVITAAHCIAERN-IVSTLNTAGEYDLSQTDPGEQTLTIE 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 VSLRYDGAHLCCGSGLLSGDWVLTAAHCFPERNRVLSRWRFAGA--VAQASPHGLQLGVQ 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 TVIIHPHF-----STKKPMYDIALKMGAFQFCHFGVGPICLPRLRQFQFAGICTTAG 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 236 AVYHGGYLPFRDPNSSENSNDIALVHLSPLPLTEYIQVCLPAAGQALVDGKICTVTG 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 185 WGRLTEGVLISQVLOEVNLPILTWEECVAAALLTLKRPISGKTFICTGFPDGRACQGD 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 296 WGNTRYTGQAGVQLQEARVPIISDVCNGADF-YGNOIKPKMF-CAGYPEGIDACQGD 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 GGSWMCR--NKGCAWTLAGVTSNGLCGRGGRNVRKSDGSGPIETDISKVLSWTHEH 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 354 GGPVCEDSISRTPRWRLCGIVSMGTGCAIAQK-----PGVYTKVSDFREWIFQA 403
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 IQT 304
DB 404 IKT 406

RESULT 10
ENTK_BOVIN
ID ENTK_BOVIN STANDARD; PRT; 1035 AA.
AC P98072;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enteropeptidase precursor (EC 3.4.21.9) (Enterokinase).
GN PRSS7 OR ENTK.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Duoenum;
RX MEDLINE=94329561; PubMed=8052624;
RA Lavallie E.R., Rehentulla A., Racie L.A., Diblasio E.A.,
RA Ferenz C., Grant K.L., Light A., McCoy J.M.;
RT "Enterokinase, the initiator of intestinal digestion, is a mosaic
RT protease composed of a distinctive assortment of domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7588-7592(1994).
RN [2]
RP SEQUENCE OF 801-1035 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=94043122; PubMed=8228855;
RA Lavallie E.R., Rehentulla A., Racie L.A., Diblasio E.A.,
RA Ferenz C., Grant K.L., Light A., McCoy J.M.;
RT "Cloning and functional expression of a cDNA encoding the catalytic
RT subunit of bovine enterokinase.";
RL J. Biol. Chem. 268:23311-23317(1993).
RN [3]
RP SEQUENCE OF 801-827.
RC TISSUE=Intestine;
RX MEDLINE=92189715; PubMed=1799406;
RA Light A., Janska H.;
RT "The amino-terminal sequence of the catalytic subunit of bovine
RT enterokinase.";
RL J. Protein Chem. 10:475-480(1991).
CC -1- FUNCTION: RESPONSIBLE FOR INITIATING ACTIVATION OF PANCREATIC
CC PROTEOLYTIC PROENZYMES (TRYPSIN, CHYMOTRYPSIN AND CARBOXYPEPTIDASE
CC A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN
CC TURN ACTIVATES OTHER PROENZYMES INCLUDING CHYMOTRYPSINOGEN,
CC PROCARBOXYPEPTIDASES, AND PROELASTASES.
CC -1- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-Ile-7 bond in
CC trypsinogen.
CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC (LIGHT) CHAIN AND A
CC MULTIDOMAIN (HEAVY) CHAIN LINKED BY A DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -1- TISSUE SPECIFICITY: INTESTINAL BRUSH BORDER.
CC -1- PTM: THE CHAINS ARE DERIVED FROM A SINGLE PRECURSOR THAT IS
CC CLEAVED BY A TRYPSIN-LIKE PROTEASE.

```

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 SEA DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U09859; AAB40026.1; -;
 CC EMBL: L19663; AAB16035.1; -;
 CC PIR: A61436; A61436.
 CC HSP: P00763; IDPO.
 CC MEROPS: S01.156; -;
 CC InterPro: IPR000859; CUB_domain.
 CC InterPro: IPR001314; Chymotrypsin.
 CC InterPro: IPR002172; LDL_recept_A.
 CC InterPro: IPR000998; MAM_domain.
 CC InterPro: IPR000082; SEA_domain.
 CC InterPro: IPR001254; Ser_protease_Try.
 CC InterPro: IPR001190; Srcr_receptor.
 CC Pfam: PF00057; ldl_recept_a; 2.
 CC Pfam: PF00089; trypsin; 1.
 CC Pfam: PF00431; CUB; 2.
 CC Pfam: PF00530; SRCR; 1.
 CC Pfam: PF00629; MAM; 1.
 CC Pfam: PF01390; SEA; 1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00042; CUB; 2.
 CC SMART: SM00192; LDLA; 2.
 CC SMART: SM00137; MAM; 1.
 CC SMART: SM00200; SEA; 1.
 CC SMART: SM00202; SR; 1.
 CC SMART: SM00020; Tryp_SPC; 1.
 CC PROSITE: PS01180; CUB; 2.
 CC PROSITE: PS01209; LDLA_1; 2.
 CC PROSITE: PS00068; LDLA_2; 2.
 CC PROSITE: PS00740; MAM_1; 1.
 CC PROSITE: PS00060; MAM_2; 1.
 CC PROSITE: PS00024; SEA; 1.
 CC PROSITE: PS00420; SRCR_1; FALSE_NEG.
 CC PROSITE: PS02087; SRCR_2; 1.
 CC PROSITE: PS02040; TRYPSIN_DOM; 1.
 CC PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC PROSITE: PS00135; TRYPSIN_SER; 1.
 CC Signal-anchors: Glycoprotein; Myristate; Hydrolase;
 CC Serine protease; Zymogen; Transmembrane; Repeat; Alternative splicing.
 CC Signal-anchors: Glycoprotein; Myristate; Hydrolase;
 CC Serine protease; Zymogen; Transmembrane; Repeat; Alternative splicing.
 CC CHAIN 1 800 NON-CATALYTIC CHAIN (HEAVY CHAIN).
 CC CHAIN 801 1035 CATALYTIC CHAIN (LIGHT CHAIN).
 CC DOMAIN 1 18 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC TRANSMEM 19 47 (POTENTIAL).
 CC DOMAIN 48 1035 EXTRACELLULAR (POTENTIAL).
 CC DOMAIN 54 169 SEA.
 CC DOMAIN 197 238 LDL-RECEPTOR CLASS A 1.
 CC DOMAIN 240 350 CUB 1.
 CC DOMAIN 358 520 MAM.
 CC DOMAIN 540 650 CUB 2.
 CC DOMAIN 657 695 LDL-RECEPTOR CLASS A 2.
 CC DOMAIN 694 787 SRCR.
 CC DOMAIN 801 1035 SERINE PROTEASE.
 CC ACT_SITE 841 841 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC ACT_SITE 892 892 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC ACT_SITE 987 987 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC LIPID 2 2 MYRISTATE (POTENTIAL).
 CC DISULFID 199 212 BY SIMILARITY.
 CC DISULFID 206 225 BY SIMILARITY.

FT DISULFID 219 236 BY SIMILARITY.
 FT DISULFID 659 671 BY SIMILARITY.
 FT DISULFID 666 684 BY SIMILARITY.
 FT DISULFID 678 693 BY SIMILARITY.
 FT DISULFID 788 812 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 826 842 BY SIMILARITY.
 FT DISULFID 926 993 BY SIMILARITY.
 FT DISULFID 957 972 BY SIMILARITY.
 FT CARBOHYD 983 1011 BY SIMILARITY.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 519 519 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 722 722 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 741 741 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 762 762 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 864 864 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 903 903 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 166 192 MISSING (IN SHORT ISOFORM).
 FT CONFLICT 808 808 R -> Y (IN REF. 3).
 SQ SEQUENCE 1035 AA; 114887 MW; E207970B08296E13 CRC64;
 Query Match 27.0%; Score 442.5; DB 1; Length 1035;
 Best Local Similarity 34.9%; Pred. No. 2.7e-33;
 Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;
 QY 36 SCGQSLV--KVOPMNYFNIESRILGSOVEKSYPMOVSLSKQKQKHIGSGIVSPQWVIT 93
 DB 787 SCGKKLVTVQEVSP-----KIVGSDSREGANFWVALFDDQOVCGASLSRDLWS 838
 QY 94 AAHCIAANRNI-VSTLNVTAGEDYLSQ-TDPGQTTITETVTHPHFSFKKPMYDIALK 151
 DB 839 AAHCYVYGRNMEPSKWKAVLGLHMASNLSPQIETRLIDQIVINPHYN-KRRKNNDIAMMH 897
 QY 152 MAGAFQFGHEVGPICLPRLREOFEAGFICTTAGWGRLEGGVLSOVLOEVNLPILTWEBC 211
 DB 898 LEMKYNVTDYIQPICLPEENQVFPFGRICISAGWALYQGSTADVLGQADVPLLSNEKC 957
 QY 212 VAAL-----LTLKRPISSGKTFCTGPDGGRDACQDGGSLMCRNKKGAWTLGAVTSWGL 267
 DB 958 QQQMPEYNIT-----ENMVCAGYEAGVDSCQDGGPLMCQ-ENNRWLLAGVTSFGY 1009
 QY 268 GCGRWVRNVRKSDGSGFGIFTDISKVLWSHEHI 302
 DB 1010 QCALPNR-----ECVYARVPRFTEWIOSFL 1034
 RESULT 11
 HEPS_MOUSE
 ID HEPS_MOUSE STANDARD; PRT; 416 AA.
 AC O35453;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serine protease hepsin (EC 3.4.21.-).
 GN HPN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.


```
Query Match          26.8%; Score 438.5; DB 1; Length 416;
Best Local Similarity 33.0%; Pred. No. 2.le-33;
Matches 102; Conservative 48; Mismatches 112; Indels 47; Gaps 11;

QY 11 KLILLGLVFFERGSAAISLPKAPSCGSLVKVQPNWYFISRLGSGQVEKGSYPWQ 70
DB 129 RLDDVISVCDPRGFLTAT---CODGRRKLPV-----DRVGGQSSLGWRPWQ 176
QY 71 VSLKQRKHGGGSIQVQWVITAHCIAANRN-IVSTLNVTAGEDYDLSDTDPGEOTLTIE 129
DB 177 VSLRDGTHLCGSLSGDWLTAACFPERNRLSRWRVFAGA--VARTSPHVAQLGVQ 234
QY 130 TVIIHPH-----STKKPMVDYDIALKMGAFQFGHFVGPICLPDLRPFQFAGFICTTAG 184
DB 235 AVIYHGGVLPFRDPTIDENSNDIALVHLSSSLPLTEYIQVCLPAAGQALVDGKVCVTG 294
QY 185 WGRLEGGVLSQVLOEVNLPILTWECEVAALLTLKRP-----ISGKFLCTGPDGGRD 238
DB 295 WNTGYFQQAQVQLQEARVPIISNEVC-----NSPDFYGNQIKPKMF-CAGYPEGGID 346
QY 239 ACQSDSGSLMCRNK---KGAWTLAGVTSWGLGCGRGNRNVRKSDQSGPFIETDISKVL 295
DB 347 ACQSDSGGHFVCEDRISGTSRRLCGIVSWGTCAL-----ARK-----PGVYTKVIDFR 396
QY 296 SWIHEHIQT 304
DB 397 EWIFQAIKT 405

RESULT 13
ENTK_PIG
ID ENTK_PIG STANDARD; PRT; 1034 AA.
AC P98074;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enteropeptidase precursor (EC 3.4.21.9) (Enterokinase).
GN PRS7 OR ENTK.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Duodenal mucosa;
RX MEDLINE=94327548; PubMed=8051081;
RA Matsushima M., Ichinose M., Yahagi N., Kakei N., Tsukada S.,
RA Miki K., Kurokawa K., Tashiro K., Shiokawa K., Shinomiya K.,
RA Uneyama H., Inoue H., Takahashi T., Takahashi K.;
RT "Structural characterization of porcine enteropeptidase.";
RL J. Biol. Chem. 269:19976-19982(1994).
CC -!- FUNCTION: RESPONSIBLE FOR INITIATING ACTIVATION OF PANCREATIC
CC PROTEOLYTIC PROENZYMES (TRYPSIN, CHYMOTRYPSIN AND CARBOXYPEPTIDASE
CC A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN
CC TURN ACTIVATES OTHER PROENZYMES INCLUDING CHYMOTRYPSINOGEN,
CC PROCARBOXYPEPTIDASES, AND PROELASTASES.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-Ile-7 bond in
CC trypsinogen.
CC -!- SUBUNIT: HETEROTRIMER OF A CATALYTIC (LIGHT) CHAIN, A MULTIDOMAIN
CC (HEAVY) CHAIN, AND A MINI CHAIN.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -!- PTM: THE CHAINS ARE DERIVED FROM A SINGLE PRECURSOR THAT IS
CC CLEAVED BY A TRYPSIN-LIKE PROTEASE.
CC -!- PTM: THE MINI CHAIN MAY BE CLEAVED BY ELASTASE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SEA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SRC DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D30799; BAA06459.1; -.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.156; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR000998; MAM_domain.
DR InterPro; IPR000082; SEA_domain.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Socr_receptor.
DR Pfam; PF00057; ldl_recept_a; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 1.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF01390; SEA; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 2.
DR SMART; SM00137; MAM; 1.
DR SMART; SM00200; SEA; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00202; Tryp_SPC; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PS00068; LDLRA_2; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00060; MAM_2; 1.
DR PROSITE; PS50024; SEA; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS50287; SRCR_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Signal-anchor; Glycoprotein; Myristate; Hydrolase;
KW Serine protease; Zymogen; Transmembrane; Repeat.
FT CHAIN 52 117 NON-CATALYTIC M CHAIN (MINI CHAIN).
FT CHAIN 118 799 NON-CATALYTIC H CHAIN (HEAVY CHAIN).
FT CHAIN 800 1034 CATALYTIC L CHAIN (LIGHT CHAIN).
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 47 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 48 1034 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 52 169 SEA.
FT DOMAIN 197 238 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 240 349 CUB 1.
FT DOMAIN 357 519 MAM.
FT DOMAIN 539 649 CUB 2.
FT DOMAIN 656 694 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 693 786 SRCR.
FT ACT_SITE 840 840 SERINE PROTEASE.
FT ACT_SITE 891 891 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 986 986 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT LIPID 2 2 MYRISTATE (POTENTIAL).
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 206 225 BY SIMILARITY.
FT DISULFID 219 236 BY SIMILARITY.
FT DISULFID 658 670 BY SIMILARITY.
FT DISULFID 665 683 BY SIMILARITY.
FT DISULFID 677 692 BY SIMILARITY.
FT DISULFID 787 911 INTERCHAIN (BY SIMILARITY).
FT DISULFID 825 841 BY SIMILARITY.
FT DISULFID 925 992 BY SIMILARITY.
FT DISULFID 956 971 BY SIMILARITY.
FT DISULFID 982 1010 BY SIMILARITY.
```


[illegible]

Db 390 CLQNPAPKELTVVLGDRHNQCEQCOTLAVRDYRLHEAFS---PITYQHDIALVRLQE 446
Qy 155 AFQ--FGH---FVCPICLPE--LREQFEAGFICTTAGWRLTEGGVLSQVLQEVNLPILT 207
Db 447 SADGCCAHPSPFVQVCLPSTAARPAESEAACEVAGWGHQFEGGEYSFLEQAQVPLID 506
Qy 208 WEECVAAALLTKRPISGKTF---LCTGFPDGGRDACQDGGSLMCRNK--KGAWTLAG 261
Db 507 PORCSAP-----DVHGAFTQGMCLCAGFLEGGTDACQDGGSLVCEDETPERQLIRG 560
Qy 262 VTSWGLCGGRGNRNVRKSDQSGFGIFTDISKVLWIHEH 301
Db 561 IVSWGSGCG----NRLK-----PGVYTDVANYLAWIREH 590

Search completed: March 25, 2003, 03:58:57
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:02:44 ; Search time 62 Seconds
(without alignments)

1016.942 Million cell updates/sec

Title: US-09-735-713a-2

Perfect score: 1636

Sequence: 1 MSLKMLISRNKLLLLGIVF.....IFTDISKVLWIHEIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	688.5	42.1	1004	13 P79953	P79953 xenopus lae
2	657	40.2	974	13 Q90WD8	Q90WD8 bufo japoni
3	626	38.3	1524	13 Q91674	Q91674 xenopus lae
4	496	30.3	277	5 Q96899	Q96899 scolopendra
5	486	29.7	787	5 Q9VEY6	Q9VEY6 drosophila
6	480.5	29.4	799	11 Q9DBI0	Q9DBI0 mus musculus
7	475.5	29.1	335	11 Q8VIF2	Q8VIF2 mus musculus
8	475.5	29.1	624	11 Q9DAT3	Q9DAT3 mus musculus
9	469.5	28.7	624	11 Q91Y47	Q91Y47 mus musculus
10	462.5	28.3	467	5 Q967X8	Q967X8 panulirus a
11	462	28.2	453	11 Q8VDE0	Q8VDE0 mus musculus
12	458	28.0	638	11 Q8E0P5	Q8E0P5 mus musculus
13	450	27.5	339	11 Q99L44	Q99L44 mus musculus
14	446	27.3	643	6 Q97506	Q97506 sus scrofa
15	440.5	26.9	433	13 Q90YK1	Q90YK1 brachydanio
16	436.5	26.7	317	13 Q9DGR3	Q9DGR3 xenopus lae

17	435.5	26.6	310	11 Q91XC4	Q91XC4 mus musculus
18	430.5	26.3	310	11 Q9QY29	Q9QY29 mus musculus
19	430	26.3	767	13 Q9DGR2	Q9DGR2 xenopus lae
20	428.5	26.2	581	5 Q96015	Q96015 drosophila
21	428.5	26.2	1047	5 Q9VZH2	Q9VZH2 drosophila
22	427.5	26.1	387	5 Q9XY57	Q9XY57 ctenocephal
23	426	26.0	314	5 Q9VR15	Q9VR15 drosophila
24	426	26.0	624	6 Q95ME7	Q95ME7 oryctolagus
25	423.5	25.9	325	5 Q15944	Q15944 sarcophaga
26	423.5	25.9	616	6 Q97507	Q97507 sus scrofa
27	423	25.9	256	5 Q9XYI1	Q9XYI1 rhyzopertha
28	422.5	25.8	267	5 Q9BK47	Q9BK47 luidia foli
29	422.5	25.8	456	6 Q9TTR0	Q9TTR0 canis famil
30	421	25.7	273	6 Q9XSM1	Q9XSM1 ovis aries
31	421	25.7	855	11 Q91J17	Q91J17 rattus norv
32	420.5	25.7	321	4 Q96R28	Q96R28 homo sapien
33	420.5	25.7	1047	5 Q24019	Q24019 drosophila
34	418.5	25.6	329	6 Q9GL10	Q9GL10 ovis aries
35	417	25.5	264	11 Q9D7P8	Q9D7P8 mus musculus
36	417	25.5	264	11 Q9ER05	Q9ER05 mus musculus
37	414.5	25.3	300	4 Q96EF3	Q96EF3 homo sapien
38	414.5	25.3	573	5 Q9V516	Q9V516 drosophila
39	414	25.3	254	5 Q8T637	Q8T637 aedes aegypt
40	414	25.3	264	11 Q9EQ28	Q9EQ28 rattus norv
41	414	25.3	268	5 Q46151	Q46151 pacifastacu
42	413	25.2	263	11 Q9DC86	Q9DC86 mus musculus
43	413	25.2	405	4 Q96E86	Q96E86 homo sapien
44	412	25.2	263	11 Q9CR35	Q9CR35 mus musculus
45	409.5	25.0	468	5 Q9U0G3	Q9U0G3 pacifastacu

ALIGNMENTS

RESULT 1

P79953	ID	P79953	PRELIMINARY;	PRT;	1004 AA.
AC	P79953;				
DT	01-MAY-1997	(TEMBLrel. 03, Created)			
DT	01-MAY-1997	(TEMBLrel. 03, Last sequence update)			
DT	01-JUN-2002	(TEMBLrel. 21, Last annotation update)			
DE	Oviductin.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;				
OC	Xenopodinae; Xenopus.				
OX	NCBI_taxid=8335;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-99184825; PubMed-10084976;				
RA	Lindsay L.L., Wieduwilt M.J., Hedrick J.L.;				
RT	"Oviductin, the Xenopus laevis oviductal protease that processes egg				
RT	envelope glycoprotein gp43, increases sperm binding to envelopes, and				
RT	is translated as part of an unusual mosaic protein composed of two				
RT	protease and several CUB domains.";				
RL	Biol. Reprod. 60:989-995(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Yang J.C., Lindsay L.L., Hedrick J.L.;				
RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.				
CC	-I- SIMILARITY: CONTAINS 3 CUB DOMAINS.				
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE				
CC	TRYPSIN FAMILY				
DR	EMBL; U81291; AAB53972.1; -				
DR	HSSP; P00763; IDPO.				
DR	MEROPS; S01.240; -				
DR	InterPro; IPR001314; Chymotrypsin.				
DR	InterPro; IPR000859; CUB domain.				
DR	InterPro; IPR001254; Ser_protease_Try.				
DR	Pfam; PF00431; CUB; 2.				
DR	Pfam; PF00089; trypsin; 2.				
DR	PRINTS; PR00722; CHYMOTRYPSIN.				
DR	SMART; SM00042; CUB; 3.				


```
QY 24 GKSAALSLPKAPSCGQSLVKVQPNYENIFSRILGGQVEKGSYPWQVSLKROKHICGG 83
Db 32 GETAELCKGTPEIGD-----EPDLETF--SRIVGGGDAVGGQPTVSLKLNHRHICGG 84
QY 84 SIVSPQWVITAAHCII--ANRNIVSTLNVNTAGEYDLSTQDGEQTLTIETVTHPHFSTKK 141
Db 85 SIVRKDMVVTAAHCVTPVTEIKSHMTVIVGEYDQVMDSQEQSIPVSHIEPHENYRGDG 144
QY 142 PMDYDIALLKAGAFQCHFGVPTCLPDLREOFAGFTCTTAGHGRLTGEGVLSQVLOEV 201
Db 145 NMGYDIALVFLSKPIIFGSGVQVQICLPQVGEKIEAGTLCVSSGWRLEENGDLSPVLOEV 204
QY 202 NLPILTWEECAALTLTKRPISGTFCTGPPDGGRCACQDGSGLMCRNKKGAWTLAG 261
Db 205 KLPVVDNCTCHAVLEIGHVPLDDTMLCAGEPEGGMACQDGSQPFVCRRRSGVWFLAG 264
QY 262 VTSNGLGCGRW--RNNVRKSDQSGPGFTDISKVLSTWHEHQTG 305
Db 265 CVSMGLGCGRWGAKIIRISGSGPAIFSRVSSVLDLFLRPPKLTG 309

RESULT 4
O96899 ID O96899 PRELIMINARY; PRT; 277 AA.
AC O96899;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE Plasmidogen activator spa.
OS Scolopendra subspinoses.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Scolopendra.
OX NCBI_TaxID=55038;
RN [1]
RP SEQUENCE FROM N.A.
RA Sohn Y.-D., You W.K., Kim K.Y., Chung K.H., Park D.-H.;
RT "Cloning of a cDNA for a novel plasminogen activator from a Korean
RT centipede Scolopendra.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; J079521; AAD00320.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.122; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD-SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 277 AA; 30172 MW; EC62F80C8CB246C CRC64;

Query Match 30.38; Score 496; DB 5; Length 277;
Best Local Similarity 38.18; Pred. No. 4e-39;
Matches 112; Conservative 46; Mismatches 98; Indels 38; Gaps 12;

QY 10 NKLILLIGVIFERKSAALSILPRAPSGQSLVKVQPNYENIFSRILGGQVEKGSYPW 69
Db 2 NSFTILI-VTF-----SLAFGSRG---IKNGP--MLDEFNRIVGGEAEPGEPW 47
QY 70 QVSLK----QRKHICGSGIVSPQWVITAAHCIAANRNIVSTLNVNTAGEYDLSTQDPPGQT 125
Db 48 QISLQVSWYSGYHYCGSGILDESNTVTAACHVEGMN-PSDLRLILAGEHNKKEDGTQW 106
QY 126 LTETVTHPHFSTKKPMYDIALLKAGAFQF-GHFVGPICLPDLRELQFPAFGFTCTTAG 184
Db 107 QVDVIDIIMHKDY-VYSTLENDIALKLAEPDLDTPTAVGSCICLPSQNNQFSGH-CIVTG 164
QY 185 WGRLTGEGVLSQVLOEVNLPILTWECAALLTLKRPISGKTFCTGPPDGGRCACQDGS 244
```

```
Db 165 WGSVREGGNSPNILOKVSVPLMTDEECSEYINIV-----DTMLCAGYAEGKDACQDGS 218
QY 245 GGSMLCRNKKGAWTLAGVTSGWGLCGGRWNRNVRKSDQSGPGIFTDISKVLSTWI 298
Db 219 GGPLVCPNGDGTYSLAGIVSWGICGAOP-RN-----PGVYTVQVSKFLDWI 262

RESULT 5
O9VEY6 ID O9VEY6 PRELIMINARY; PRT; 787 AA.
AC O9VEY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SB gene product.
GN SB OR CG4316.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatiades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.M., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.M.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AE003712; AAF55277.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.225; -.
DR FlyBase; FBgn0003319; Sb.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002965; P_rich_extensin.
DR InterPro; IPR001254; Ser_protease_Try.
```

DR	Pfam:	PF00089;	trypsin; 1.
DR	PRINTS:	PR00722;	CHYMOTRYPSIN.
DR	PRINTS:	PRO1217;	PRICHTEXTENS.
DR	SMART:	SM00020;	Tryp_SPC; 1.
DR	PROSITE:	PS00240;	TRYPSIN_DOM; 1.
DR	PROSITE:	PS00135;	TRYPSIN_SER; 1.
KW	Hydrolase:	Serine protease.	
SQ	SEQUENCE	787 AA;	85143 MW; 02B2B8758BD6025A CRC64;

Query Match	29.7%;	Score 486;	DB 5;	Length 787;
Best Local Similarity	37.5%;	Pred. No. 1.4e-37;		
Matches	98;	Conservative	50;	Mismatches 87; Indels 26; Gaps 18;

Qy	54	SRIILGGSQVEKGSYPQVMSLKQ-----RKHICGSGSVSPQWVITAAHCTANRNIVSTL	107
Dd	542	TRIVGGSAAPGRWPQVSVARTSFEGFSFTHRCGGALINENWIATAGHCVDLDL-LISQI	600
Qy	108	NVTAGEYDLSQTDPGCEQTLTIE----TVIIHPHFSTRKKPMYDIALLKMAGAFQFGHFVG	163
Dd	601	RIRVGEYDFSHVQ--EQLPYIERGVAKKVHPKYSE-LTYEYDLALVKLEQPLEFAPHVS	657
Qy	164	PICLPRLREQFEAFGICTTAGWGRLTEGGVLSVLQOEVLNLFILLTWEECVAALLTL-KRPI	222
Dd	658	PICLPLE-TDSLLIGMNAVTVGWGLSEGGLTPSLQEVSVPIDNVNDCKFMFRAGRQEF	716
Qy	223	SGKTFLCTGFPDGGRDACQDGSGLCRNKKGAWTLAGVTSWGLCGRGWRNNVNRKGDQ	282
Dd	717	IPDFLCAGYETGQDSCQDSSGGPQAOKSDQGRFFLAGIISWGIGC-----AEA	766
Qy	283	GSPGIFTDISKVLWSIHFIHQ	303
Dd	767	NLPGVCTRISKFTPTWILEHVR	787

RESULT 6			
Q9DBIO	ID	Q9DBIO	PRELIMINARY; PRT; 799 AA.
AC	Q9DBIO;		
CD	Q9DBIO;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBurel. 17, Last sequence update)		
DE	01-JUN-2002 (TrEMBurel. 21, Last annotation update)		
DE	1300008A22Rik protein.		
GN	1300008A22Rik.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
NCBI_TaxId	=10090;		
RP	[1]		
RR	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=LIVER;		
RX	MEDLINE=21055660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo U., Nikolaide I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	L Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nardone P., Ring B., Schenwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RA	Hayashizaki Y.;		
RA	Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
CC	-1- SIMILARITY: CONTAINS 1 CUB DOMAIN.		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE		

Query Match 29.1%; Score 475.5; DB 11; Length 624;
Best Local Similarity 35.1%; pred. No. 1e-36;
Matches 99; Conservative 47; Mismatches 113; Indels 23; Gaps

Qy	83	GSIVSPQWVITAAHCIAANRNVSTLVNTAGEYDLSDOPDGEQTLTITETVIIHPHFSTKKP	1421
Db	417	GSIIIGNOWILTAACHFSGIETPKKRLRYGGIVGNQSEINECTAFFRVOEMIIHQYTTAE-	4751
Qy	143	MDYDIALLKWAGAFQFGHVGVPICLPRLERQFEAGFICTTAGWGRLTGEGVLSQVLQEVN	1021
Db	476	SGYDIALLKLESAMNYTDFQRPICLPSKGRNVAHTECWYTGWGTALRGVEVSTLOKAK	5351
Qy	203	LPILTWEECVAAALLTLKRPISGKTFCTGPPDGGRDACQDGSGLMCRNKKKGAWTLAGV	2621
Db	536	VPLVSNSECOTRY--RRHKITNK-MICAGYKEGGKDTCKDGSGLSCK-YNGVWHLVGI	5911
Qy	263	TSWGLGCGRGNRNVKSDGSGPFIPTDISKVLSWIHEHIQT	304
Db	592	TSWGEGCGQKER-----PCVYTNVAKYVDWILEKTQT	623
RESULT 9			
Q91Y47	PRELIMINARY; PRS; 624 AA.		
ID	Q91Y47		
AC	Q91Y47;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Coagulation factor XI.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=B10.WR; TISSUE=LIVER;		
RX	MEDLINE=97385041; PubMed=9242536;		
RA	Gallani D., Sun M.F., Sun Y.;		
RT	"A comparison of murine and human factor XI.";		
RL	Blood 90:1055-1064(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=B10.WR; TISSUE=LIVER;		
RX	Gallani D., Sun M.F., Sun Y.;		
RA	"A comparison of murine and human factor XI.";		
RL	Blood 90:1055-1064(1997).		

```
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF356627; AAK40233.1; -.
DR MEROPS; S01.213; -.
DR InterPro: IPR00177; Apple.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE; PS00495; APPLE; UNKNOWN_2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 624 AA; 69874 MW; 49D281BFAEC12A03 CRC64;

Query Match      28.7%; Score 469.5; DB 11; Length 624;
Best Local Similarity 35.1%; Pred. No. 3.9e-36;
Matches 99; Conservative 46; Mismatches 114; Indels 23; Gaps 6;

QY 23 RGSALSLPKAPSCQSLVKVOPWNYNIFSRILGSGQVEKGYPMQVSLKQKQKHCIG 82
DB 365 RGLSGYSLRLCKMDNVCTTKINP-----RVVGGAAASVHGEPWQVTLHISQGHLCG 146

QY 83 GSTVSPQWVITAAHCIAANRNVSTLNVTAGEDVLSQTDPEQTLTIETVLIHPHFSTKKP 142
DB 417 GSIIGNQWILTAHCFSGIETPKLVYGVINQSEINSGTAFRDEMIHDDQYTTAE- 475

QY 143 MDYDIALKMGAFQFHFVGPICLPPELREQFEAGICTAGWGRLTGEGVLSQVLEQYN 202
DB 476 SGYDIALKLESAMNYTDFQRPICLPKGRNAVHTECVTWGTYALRGVQSTLQKAK 535

QY 203 LPILTWECEVAALLTKRPIISKTFCTGFPDGGDRDACQDGGSGGLMCRNKKGAWTLAV 262
DB 536 VPLVSNEECQTRY--RRHITNK-MTCAGYKEGGKDKTCKGDSGGPLSCK-YNGVHLVGI 591

QY 263 TSWGLCGGGRNVRNKRKSDQSGPIFTDISKVLWHEHIQT 304
DB 592 TSWGECGGQKER-----PGVTINAKYVDWILEKTQT 623

RESULT 10
Q967X8 PRELIMINARY; PRT; 467 AA.
AC Q967X8:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CUB-serine protease.
OS Panulirus argus (Spiny lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;
OC Palinuroidea; Palinuridae; Panuliridae; Panulirus.
OX NCBI_TaxID=6737;
RN [1]
RP SEQUENCE FROM N.A.
RA Levine M.Z., Walshall W.W., Tai P.C., Derby C.D.;
RT "Molecular cloning, characterization, cellular localization and
RT possible function of a CUB-serine protease in the olfactory system of
RT the spiny lobster Panulirus argus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF357226; AAK48894.1; -.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00431; CUB; 1.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 467 AA; 50453 MW; 1D2E663D3314BBFD CRC64;

Query Match      28.3%; Score 462.5; DB 5; Length 467;
Best Local Similarity 35.1%; Pred. No. 1.3e-35;
Matches 102; Conservative 57; Mismatches 115; Indels 34; Gaps 8;
```

```
Best Local Similarity 38.2%; Pred. No. 1.2e-35;
Matches 105; Conservative 40; Mismatches 97; Indels 33; Gaps 9;

QY 35 PSCGQSLVQVPMWNYNIFSRILGSGQVEKGYPMQVSLKQKQKHC-ICGSGIVSPQWVIT 93
DB 216 PSCACGNV-----NRATRVGGQETEVNEYPMQVLLVTRDMYVICGSIISQWVLT 267

QY 94 AAHCIAANRNVSTLNVTAGEDVLSQTDPEQTLTIETV--IIHPHFSTKKPMDYDIALLK 151
DB 268 AAHCVDGGNGIVL---VGDHNFASDDTTTSLRLEVVOIISHPDYDS-STVDNDNALLR 323

QY 152 MAGAFQFHFVGPICLPPELREQFEAGICTAGWGRLTGEGVLSQVLEQYNLPILTWECE 211
DB 324 LGEALEFTREVAPVCLPSNPTEDIAGVTATVTGAGATTEGGSMVTVLQEVDPVLTAAAC 383

QY 212 VAAALTLLKRPISGKTFCTGFPDGGDRDACQDGGSLMCRNKKGAWTLAVTWSWGLGCGR 271
DB 384 SWSYSLT-----ANMCAAGFSNEGDCQDGGSGGP-MVYSATSNIEQIGVSVWGRGCAR 437

QY 272 GWRNVRNKRKSDQSGPIFTDISKVLWHEHIQTGN 306
DB 438 -----PGFPGVYARVTEVLEWIAAN--TGN 460

RESULT 11
Q8VDE0 PRELIMINARY; PRT; 453 AA.
AC Q8VDE0:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE TMPRSS3 protein.
GN TMPRSS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Guipponi M., Scamuffa N., Scott H.S., Rossier C., Antonarakis S.E.;
RT "Isolation and characterization of the mouse Tmprs3 gene.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ300738; CAC83350.1; -.
DR MGD; MGI:2155445; Tmprs3.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001190; Sctr_receptor.
DR Pfam: PF00057; ldl_recept_a; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; UNKNOWN_1.
DR PROSITE; PS01068; LDLRA_2; 1.
DR PROSITE; PS0287; SSCR_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ SEQUENCE 453 AA; 49529 MW; 21E5697DC8781BD3 CRC64;

Query Match      28.2%; Score 462; DB 11; Length 453;
Best Local Similarity 33.1%; Pred. No. 1.3e-35;
Matches 102; Conservative 57; Mismatches 115; Indels 34; Gaps 8;

QY 1 MSKMLISRKNKILLIGVIFFERG-KSALSLPKAPSCQSLVKVQPMWNYNIFSRILG 59
DB 170 VSNHLLSDDKVYTLHSHSVYMRGCTSGHVTLKCSACGTRT-----CYSPIVGG 220

QY 60 SOVEKGSYPWQVSLKQKQKHCIGSGIVSPQWVITAAHCIAANRNVSTLNVTAGEDVLS 119
DB 221 NMSSTQPMQVSLQFGYHLGCGSIITPLTWITAAHCYDLYHPKSWTVQVGL--VSLM 278
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 25, 2003, 01:17:33 ; Search time 1806 seconds
(without alignments)
2744.086 Million cell updates/sec

Title: US-09-735-713A-2
Perfect score: 1636
Sequence: 1 MSLKMLISRNKLLLLGIVF.....IFTDISKVLSTWIEHIQTGN 306

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q/cgcn2_1/USPTO_spool/US09735713/runat_18032003_124142_21598/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=fst -MINMATCH=0.1 -LOPCPU=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPEXT=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713 -CGCN_1.1_1906 -runat_18032003_124142_21598 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	968	59.2	670	10	BB625475	BB625475 Mus muscu
2	480.5	29.4	3030	11	AK004939	AK004939 Mus muscu
3	475.5	29.1	2246	11	AK005546	AK005546 Mus muscu
4	426	26.0	3248	11	BC008514	BC008514 Mus muscu
5	424.5	25.9	1052	13	BI554641	BI554641 603235988
6	422.5	25.8	999	9	AL551470	AL551470 AL551470
7	420	25.7	1188	11	BC034294	BC034294 Homo sapi
8	418.5	25.6	790	13	BJ509337	BJ509337 BJ509337
9	417	25.5	868	11	AK009019	AK009019 Mus muscu
10	415.5	25.4	1056	14	BM919200	BM919200 AGENCOURT
11	415.5	25.4	1323	11	AK014645	AK014645 Mus muscu
12	414	25.3	1629	11	AK010640	AK010640 Mus muscu
13	413	25.2	885	11	AK003060	AK003060 Mus muscu
14	412.5	25.2	921	13	BI833975	BI833975 603085088
15	412	25.2	876	11	AK007765	AK007765 Mus muscu
16	412	25.2	877	11	AK003079	AK003079 Mus muscu
17	412	25.2	877	11	AK007815	AK007815 Mus muscu
18	412	25.2	877	11	AK008729	AK008729 Mus muscu
19	412	25.2	877	11	AK008927	AK008927 Mus muscu
20	412	25.2	879	11	AK008888	AK008888 Mus muscu
21	412	25.2	943	14	BQ222778	BQ222778 AGENCOURT
22	411.5	25.2	807	13	BI834199	BI834199 603084101
23	410	25.1	870	11	AK007333	AK007333 Mus muscu
24	407	24.9	1050	11	AK006271	AK006271 Mus muscu
25	406	24.8	876	11	AK007566	AK007566 Mus muscu
26	405.5	24.8	905	14	BQ729181	BQ729181 AGENCOURT
27	403	24.6	870	14	BM919892	BM919892 AGENCOURT
28	400.5	24.5	855	14	BQ956801	BQ956801 AGENCOURT
29	399.5	24.4	671	13	BI771604	BI771604 603058729
30	399	24.4	890	11	AK008644	AK008644 Mus muscu
31	399	24.4	1085	14	BM919045	BM919045 AGENCOURT
32	398.5	24.4	780	13	BI838552	BI838552 603086213
33	396.5	24.2	1115	14	BM918560	BM918560 AGENCOURT
34	393.5	24.1	876	9	AL555870	AL555870 AL555870
35	391.5	23.9	1024	11	AK006746	AK006746 Mus muscu
36	390.5	23.9	936	9	AL578261	AL578261 AL578261
37	390.5	23.9	1814	11	AK002694	AK002694 Mus muscu
38	390	23.8	896	11	AK009129	AK009129 Mus muscu
39	390	23.8	897	11	AK010149	AK010149 Mus muscu
40	387	23.7	668	13	BM641281	BM641281 170006873
41	386	23.6	695	13	BM576074	BM576074 170006871
42	386	23.6	838	13	BI834360	BI834360 603084395
43	386	23.6	925	14	BQ228286	BQ228286 AGENCOURT
44	385.5	23.6	870	11	AK007773	AK007773 Mus muscu
45	385.5	23.6	1006	11	AK008460	AK008460 Mus muscu

ALIGNMENTS

RESULT 1
BB625475
LOCUS BB625475 RIKEN full-length enriched, adult male epididymis Mus
DEFINITION musculus CDNA clone 9230106D23 5', mRNA sequence.
ACCESSION BB625475
VERSION BB625475.1 GI:16463737
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 670)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 FEATURES Location/Qualifiers
 source
 1. .790
 /organism="Oryzias latipes"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA007022"
 /clone_lib="MF01FSA CDNA"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="fry stage 40"
 BASE COUNT 176 a 204 c 231 g 178 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2.66e-36 Length: 790
 Score: 418.50 Matches: 93
 Percent Similarity: 53.88% Conservative: 46
 Best Local Similarity: 36.05% Mismatches: 84
 Query Match: 25.58% Indels: 35
 DB: 13 Gaps: 10

US-09-735-713a-2 (1-306) x BJ509337 (1-790)

QY 54 SerArgLeuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
 DB 788 ACCAGGTGTCGGTGAAGATGTCGAGCGCACAGCTGGCCCTGGCAGGTGTCCTG 729
 QY 74 Lys-----GlnArgGlnLysHisIleCysGlyGlySerIleValSerProGln 89
 DB 728 CAGTACGTAGTGTAACACTTATTACCACACCTGTGGGGTACCCTGATCTCCACAGG 669
 QY 90 TrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnVal 109
 DB 668 TGGGTCTCTACTGCTGCTACTGCTCATCGTGAAGTCG-----ACGTACAGAGTG 621
 QY 110 ThrAlaGlyGluTyrAspLeuSer---GlnThrAspProGlyGluGlnThrLeuThrIle 128
 DB 620 TACATCGGAACACAGACCTGAGCGCCACACAGAGCTGGTCCATCGCCATCGGCCCC 561
 QY 129 GluThrValIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAla 148
 DB 560 CAAAGGATCATCTCCACAGAGAAGCTGGGACTTTACAC---ATCAGAAATGACATTGCC 504
 QY 149 LeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuPro 168
 DB 503 CTGATCAAACTGTCAACTCCAGTCACCTTTTCCAAACACCACTTGGGCTGCTGCTTCCA 444
 QY 169 GluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeu 188
 DB 443 AAGCTGTGTAATCTCGGATGAGGTCGTCCTCTGATGTACCGGCTGGGACGTCCTC 384
 QY 189 ThrGluGlyGlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrp 208
 DB 383 TGGACTGGAGTCCCATGTGATATCTCTGACAGCCCTCTCCAGTGGTGGACAC 324
 QY 209 GluGluCysValAlaAlaLeuLeuThrLeuLysArgPro-----IleSer 223
 DB 323 TCCACCTGC-----AGCAGGCTGCTGCTGGTGGGCACTATCGTC 285
 QY 224 GlyLysThrPheLeuCysThrGlyPheProAspGlyArgAspAlaCysGlnGlyAsp 243
 DB 284 ACCACCAACATGATTTGTCGCGGA---GGAGCGGAGAGCTGGCTAGCTCAACGGAGAC 228
 QY 244 SerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThr 263

Db 227 TCTGGTGGCCCTGAACCTGTGCGAACCCCTGATGGCTCTCTGGGATGTCCACGGAATGTG 168
 QY 264 SerTrpGly-----LeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAsp 281
 Db 167 AGCTCGGCTCCACGCTGGGTGC-----AACTACCCCAAGAAG----- 129
 QY 282 GlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHis 299
 Db 128 -----CCCTCTGCTTCACCAGAGTACGGCGGTACATCAGCTGAGTCAAC 84
 RESULT 9
 AK009019 868 bp mRNA linear HTC 19-JAN-2002
 LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched
 DEFINITION library, clone:230003F21:chymotrypsin-like, full insert sequence.
 ACCESSION AK009019
 VERSION AK009019.1 GI:12843558
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
 clone_lib:RIKEN full-length enriched mouse cDNA library
 clone:230003F21.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20493374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Alzawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
 Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
 Alzawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,I.,
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
 Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
 Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
 Kuehl,P., Lewis,S., Matsuo,Y., Nikaudo,I., Pesole,G.,
 Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
 Baldarelli,R., Barsh,G., Boffelli,D., Bult,C.,
 Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
 Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
 Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
 Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
 Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
 Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
 Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
 Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J., Kohtsuki,S.
 and Hayashizaki,Y.
 TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Nature 409 (6921), 685-690 (2001)
21085660
11217851
5 (bases 1 to 868)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAAGATCAAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGATCTCGATTAAATTAATCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

COMMENT

FEATURES
source

Location/Qualifiers
1..868
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2300003F21"
/db_xref="MGD:MG1:1893096"
/db_xref="taxon:10090"
/clone="2300003F21"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..868
/gene="Ctrl"
19..813
/gene="Ctrl"
/note="chymotrypsin-like
data source:MGD, source key:MG1:88558, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB26029.1"
/db_xref="GI:12843559"
/db_xref="MGD:MG1:88558"
/translation="MLLSLTLSVLILSSWCGVPATIPALSYNQRIVNGENAVPGS
WPQVSLQDNTGFHCGGSLIPNNVYTAHQVTPGRHFVILGEYDRSSNAEPVQVL
SIARATHPNWNNMNDLTLLKLAAPRYTAQVSPVCLASTNEALPSGLTCVTGW
GRISGVNTPARLQOVLPVYNQCRWNGARITDAMICAGSGASSCGGSGGPL
VCQKNTWVLIGVSWGTKNICNIQAPMYTRKVSFSTWVNMCMN"
846..851
/gene="Ctrl"
/note="putative"
868
/gene="Ctrl"

gene

CDS

polyA_signal

polyA_site

BASE COUNT	183 a	262 c	217 g	206 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	4.5e-36	Length:	868		
Score:	417.00	Matches:	101		
Percent Similarity:	48.51%	Conservative:	46		
Best Local Similarity:	33.33%	Mismatches:	104		
Query Match:	25.49%	Indels:	52		
Db:	11	Gaps:	12		
US-09-735-713A-2 (1-306) x AK009019 (1-868)					
QY	5	MetLeuLeuSerArgAsnLysLeuLeuLeuLeuGlyValPhePheGluArgGly	24		
Db	22	CTACTGCTCAGCCTAACCCCTTAGCCCTGGTCTCCTTGGC-----	60		
QY	25	LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal	44		
Db	61	-----TCTCTCTGGGGCTGGTGGTCTTCTGCGCATC	90		
QY	45	GlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLys	64		
Db	91	ACGCTGCACTGAGCTACAAT-----CAGAGAATTTGTCAACGGGAGAAATGCATGCCA	144		
QY	65	GlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLys---HisIleCysGlyGly	83		
Db	145	GGCTCTGGCCCTGCGAGGTGCTCTCCAGGATACACCGGCTTCCACTTCTGCGGTGT	204		
QY	84	SerIleValSerProGlnTrpValIleThrAlaAlaHisCys-----IleAlaAsnArg	101		
Db	205	TCTCTCATCAGTCCGCAACTGGTGTGTCAGCGGTGCCACTGCCAAGTCACGCTGGACGC	264		
QY	102	AsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspPro	121		
Db	265	CACTTT-----GTCGTTTTGGGAGAATATGACCCGATCTTCCAATGCTGAA	309		
QY	122	GlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLys	141		
Db	310	CCTGTGCGAGGTCTCTCGATCGCAAGGGCCATCATCTACCTTACTTGGNACGCC---	366		
QY	142	ProMetAspTyrAspIleAlaLeuLeuLysMetalGlyAlaPheGlnPheHisPhe	161		
Db	367	ACCATGAACAATAGCTGACTCTCTGAAAGTTGCTCGCCAGCCCGGTACACAGCAAA	426		
QY	162	ValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThr	181		
Db	427	GTCTCACCAGTCTGCTGGCTTCCACAAACGAGGCACCTGCTTCGGGGCTCACCTGTGTC	486		
QY	182	ThrAlaGlyTrpGlyArgLeuThr---GluGlyGlyValLeuSerGlnValLeuGlnGlu	200		
Db	487	ACCCTGGCTGGGGCCGAATCAGTGGTGGGCAATGTGACACCACTCGCTGGCAGCAA	546		
QY	201	ValAsnLeuProIleLeuThrTrpGluGluCys-----ValAlaAlaLeuLeuThr	217		
Db	547	GTGTCTTACCCCTGGTCACTGTGAATCAGTGTGCGCAGTACTGGGGTGCACGCATTACC	606		
QY	218	LeuLysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyArg	237		
Db	607	-----GATCCCATGATATATGTCAGGT-----GGCTCAGCGGCC	639		
QY	238	AspAlaCysGlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaTrp	257		
Db	640	TCCATCATCAGGGTGACTCAGGAGCCCTCTTGTCTGCCAG---AAGGAAACACCTGG	696		
QY	258	ThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnVal	277		
Db	697	GTGCTTATTGGGATTGTTCTCTCTGGGGC-----ACT	726		
QY	278	ArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrp	297		
Db	727	AAGAACTCAACATACAAGACCCGCCCATGTACACTCGGTCAGTTCAGTACCTAGT	786		

```

QY 298 ileHisGlu 300
|||:||||:
Db 787 ATCAACCA 795

RESULT 10
BM919200
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM919200 1056 bp mRNA linear EST 12-MAR-2002
AGENCOURT_6715624 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5748414
5', mRNA sequence.
BM919200.1 GI:19369579
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1056)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12776 row: c column: 07
High quality sequence stop: 673.
Location/Qualifiers
1..1056
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5748414"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 214 a 323 c 273 g 217 t 29 others
ORIGIN

Alignment Scores:
Pred. No.: 8,88e-36 Length: 1056
Score: 415.50 Matches: 100
Percent Similarity: 51.09% Conservatve: 41
Best Local Similarity: 36.23% Mismatches: 96
Query Match: 25.40% Indels: 39
DB: 14 Gaps: 13

US-09-735-713a-2 (1-306)' x BM919200 (1-1056)

QY 36 SerCysGlyGlnSerLeuValGlnProTyrPheAsnIlePheSerArg 55
|||:||||:
Db 46 TCCTGGGCGTGGCGATCTCTCGCATCAACCGGCACTGAGCTTCAGC 99
|||:||||:

QY 56 IleLeuGlyGlySerGlnValGluLysGlySerTyrProTyrPheValSerLeuLysGln 75
|||:||||:
Db 100 ATTGTCACCGGGAGATGACAGTGTGGGCTCTGGGCGCTGGAGGTCCTCGGCGAGC 159
|||:||||:

QY 76 ArgGlnLys---HisIleCysGlyGlySerIleValSerProGlnTrpValIleThra 94
|||:||||:
Db 160 AGCAGCGGCTTCACACTTCTCGGGTGGTCTCTCATCAGCCAGCTCTGGGTGGTCACTGCT 219
|||:||||:

95 AlaHisCysIleAlaAsn-----ArgAsnIleValSerThrLeuAsnValThraIaGly 112
|||:||||:
Db 220 GCCCACTGCAATGTACAGCCCTGGCGCCATT-----GTTGCTCTGGGC 264
|||:||||:

QY 113 GluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrValIle 132
|||:||||:
Db 265 GAGTATGACCGCATCATCAACAGCAGAGCCCTTCAGGTTCTGCGGTCTCGGGCCATT 324
|||:||||:

QY 133 IleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLysMet 152
|||:||||:
Db 325 ACACACCTAGCTGGAAGTCT---ACCACCATGACCAATGACGTGAGCTGCTGAAGCTC 381
|||:||||:

QY 153 AlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArgGlu 172
|||:||||:
Db 382 GCCTGCCAGCCAGTACACACACAGCATCTCGCCAGTTTGCCTGGCATCTCAACAGCAG 441
|||:||||:

QY 173 GlnPheGluAlaGlyPheIleCysThrAlaGlyTrpGlyArgLeuThr---GluGly 191
|||:||||:
Db 442 GCTCTGACTGAAGCCCTCAGCTGTGTCCACCGGCTGGGCTGCGCTCAGTGGCGTGGC 501
|||:||||:

QY 192 GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTrpGluGluCys 211
|||:||||:
Db 502 AATGTGACACGACGACATCTGCAGCAGTGGCTTGGCCCTGGTCACTGTGAATCACTGC 561
|||:||||:

QY 212 ValAlaAlaLeuLeuThrLysArgProIleSerGlyLys-----ThrPhe 227
|||:||||:
Db 562 -----CGGCAGTACTGGGCTCAAGTATCACTGACTCCATG 597
|||:||||:

QY 228 LeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247
|||:||||:
Db 598 ATCTGTGAGGT-----GGCGAGGTGCTCTCGTGCAGGCTGAGTCCGGAGGCGCT 651
|||:||||:

QY 248 LeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeu 267
|||:||||:
Db 652 CTGTCTGCCAG---AAGGGAACACATGGGTGCTTATTGTGTCTCTCTGGGCGACC 708
|||:||||:

QY 268 ---GlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySerProGly 286
|||:||||:
Db 709 AAAAAGTGC-----AATGTGGC-----GCACCTGCT 735
|||:||||:

QY 287 IlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302
|||:||||:
Db 736 GTGTATCTCGAGTTAGTAGTTCAGCACCTGGATCAACACAGGTCATA 783
|||:||||:

RESULT 11
AK014645 1323 bp mRNA linear HTC 19-JAN-2002
AK014645 Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
library, clone:4733401N09:BRAIN SPECIFIC SERINE PROTEASE-4 TYPE1,
full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

```


Alignment Scores:	
Pred. No.:	2.49e-35
Score:	414.00
Length:	1629
Matches:	104
Percent Similarity:	49.6%
Best Local Similarity:	34.50%
Conservative:	44
Mismatches:	115
Query Match:	25.31%
Indels:	38
DB:	11
Gaps:	9

US-09-735-713A-2 (1-306) x AK010640 (1-1629)

12	LeuIleuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAlaLeuSerSerLeu	31	LeuIleuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAlaLeuSerSerLeu
QY	12 LeuIleuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAlaLeuSerSerLeu	QY	31 LeuIleuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAlaLeuSerSerLeu
Db	12 LeuIleuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAlaLeuSerSerLeu	Db	31 LeuIleuLeuLeuGlyIleValPhePheGluArgGlyLysSerAlaAlaLeuSerSerLeu
163	ATTCTGCTCTCTTCGAGTGTCTC-----CAGTCGGGAATCCGAGGTGAC	207	ATTCTGCTCTCTTCGAGTGTCTC-----CAGTCGGGAATCCGAGGTGAC
Db	163 ATTCTGCTCTCTTCGAGTGTCTC-----CAGTCGGGAATCCGAGGTGAC	Db	207 ATTCTGCTCTCTTCGAGTGTCTC-----CAGTCGGGAATCCGAGGTGAC
32	ProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPheAsn	51	ProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPheAsn
QY	32 ProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPheAsn	QY	51 ProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPheAsn
Db	32 ProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPheAsn	Db	51 ProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPheAsn
208	GGGACTGAAGCCCTCCTGTGCT-----GCCGTGTCATCCAGCCA-----	243	GGGACTGAAGCCCTCCTGTGCT-----GCCGTGTCATCCAGCCA-----
Db	208 GGGACTGAAGCCCTCCTGTGCT-----GCCGTGTCATCCAGCCA-----	Db	243 GGGACTGAAGCCCTCCTGTGCT-----GCCGTGTCATCCAGCCA-----
52	IlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnVal	71	IlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnVal
QY	52 IlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnVal	QY	71 IlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnVal
Db	52 IlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnVal	Db	71 IlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnVal
244	-----CGCATCACCCTGGTGGCAGTGCAGAAAGCCCGGTGAGTGGCCCTGGCAGGTC	294	-----CGCATCACCCTGGTGGCAGTGCAGAAAGCCCGGTGAGTGGCCCTGGCAGGTC
Db	244 -----CGCATCACCCTGGTGGCAGTGCAGAAAGCCCGGTGAGTGGCCCTGGCAGGTC	Db	294 -----CGCATCACCCTGGTGGCAGTGCAGAAAGCCCGGTGAGTGGCCCTGGCAGGTC
72	SerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpVal	91	SerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpVal
QY	72 SerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpVal	QY	91 SerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpVal
Db	72 SerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpVal	Db	91 SerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpVal
295	AGCATCACCTACCATGGCAACCATGTTGTGGCGGGTGGCTGTGTCAATAAATGGGTG	354	AGCATCACCTACCATGGCAACCATGTTGTGGCGGGTGGCTGTGTCAATAAATGGGTG
Db	295 AGCATCACCTACCATGGCAACCATGTTGTGGCGGGTGGCTGTGTCAATAAATGGGTG	Db	354 AGCATCACCTACCATGGCAACCATGTTGTGGCGGGTGGCTGTGTCAATAAATGGGTG
92	IleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnValThrAla	111	IleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnValThrAla
QY	92 IleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnValThrAla	QY	111 IleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnValThrAla
Db	92 IleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnValThrAla	Db	111 IleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnValThrAla
355	GTGTGCTGCTCCTCCTGCTTCCCCAGAGAACACACAGCAGGGAACGCTATGAGGTGAAGCTG	414	GTGTGCTGCTCCTCCTGCTTCCCCAGAGAACACACAGCAGGGAACGCTATGAGGTGAAGCTG
Db	355 GTGTGCTGCTCCTCCTGCTTCCCCAGAGAACACACAGCAGGGAACGCTATGAGGTGAAGCTG	Db	414 GTGTGCTGCTCCTCCTGCTTCCCCAGAGAACACACAGCAGGGAACGCTATGAGGTGAAGCTG
112	GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	131	GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal
QY	112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	QY	131 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal
Db	112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	Db	131 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal
415	GGGGCCCCACAGCTAGCTACTCTACACCAATGACACTGTGTCACACAGTGGCTCAGATC	474	GGGGCCCCACAGCTAGCTACTCTACACCAATGACACTGTGTCACACAGTGGCTCAGATC
Db	415 GGGGCCCCACAGCTAGCTACTCTACACCAATGACACTGTGTCACACAGTGGCTCAGATC	Db	474 GGGGCCCCACAGCTAGCTACTCTACACCAATGACACTGTGTCACACAGTGGCTCAGATC
132	IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys	151	IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys
QY	132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys	QY	151 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys
Db	132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys	Db	151 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys
475	ATCACCCACTCAAGCTAC---CGAAGAAGAGGGTCCCGAGGGGACATCGCTCATCCCG	531	ATCACCCACTCAAGCTAC---CGAAGAAGAGGGTCCCGAGGGGACATCGCTCATCCCG
Db	475 ATCACCCACTCAAGCTAC---CGAAGAAGAGGGTCCCGAGGGGACATCGCTCATCCCG	Db	531 ATCACCCACTCAAGCTAC---CGAAGAAGAGGGTCCCGAGGGGACATCGCTCATCCCG
152	MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171	MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg
QY	152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	QY	171 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg
Db	152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	Db	171 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg
532	CTCAGCAGTCTCTCACCCTTCTCCCGCTACATCAGACCCCATCTGCCTCCCTGCAGCCAAT	591	CTCAGCAGTCTCTCACCCTTCTCCCGCTACATCAGACCCCATCTGCCTCCCTGCAGCCAAT
Db	532 CTCAGCAGTCTCTCACCCTTCTCCCGCTACATCAGACCCCATCTGCCTCCCTGCAGCCAAT	Db	591 CTCAGCAGTCTCTCACCCTTCTCCCGCTACATCAGACCCCATCTGCCTCCCTGCAGCCAAT
172	GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	191	GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly
QY	172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	QY	191 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly
Db	172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	Db	19

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M1563 row: p column: 21
 High quality sequence stop: 815.

Location/Qualifiers
 1. 921

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5224196"
 /clone_lib="NIH_MGC_120"
 /lab_host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dr primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."
 BASE COUNT 201 a 278 c 240 g 202 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.57e-35 Length: 921
 Score: 412.50 Matches: 99
 Percent Similarity: 50.92% Conservative: 40
 Best Local Similarity: 36.26% Mismatches: 95
 Query Match: 25.21% Indels: 39
 DB: 13 Gaps: 13

US-09-735-713A-2 (1-306) x BI833975 (1-921)

Qy 36 SerCysGlyGlnSerLeuValLysGlnProTyrPheAsnIlePheSerArg 55
 Db 19 TCCTGGGCTGGCGCATCTCCATCAACCGGCACCTGAGCTTCAGC-----CAGAGG 72
 Qy 56 IleLeuGlyGlySerGlnValLysGlySerTyrProTyrPheGlnValSerLeuLysGln 75
 Db 73 ATTGTCAACGGGGAGATGTCAGTGTGGGCTCTGGCCCTGGCAGGTGCTCCCTCAGGAC 132
 Qy 76 ArgGlnLys---HisIleCysGlyGlySerIleValSerProGlnTyrValIleThrAla 94
 Db 133 AGCAGCGCTTCCACTCTCGGTGGTCTCTCATCAGCCAGTCTCGGTGGTCACTGCT 192
 Qy 95 AlaHisCysIleAlaAsn-----ArgAsnIleValSerThrLeuAsnValThrAlaGly 112
 Db 193 GCCCACTGCAATGTGACGCCCTGGCGCCCATTTT-----GTTGCTCTGGGC 237
 Qy 113 GluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrValIle 132
 Db 238 GAGTATGACCATCATCAACACGAGCCCTGTGAGGTCTGTCCGCTCTCTCGGGCCATT 297
 Qy 133 IleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLysMet 152
 Db 298 ACACACCTAGCTGGAATCT---ACCACCATGAACATGACGTGACGCTCTCAAGCTC 354
 Qy 153 AlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArgGlu 172
 Db 355 GCCTGCCGACGCCAGTACACACGATCTCGGCAGTTTGCCTGGCATCTCTCAACAG 414
 Qy 173 GlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThr---GluGly 191
 Db 415 GCTCTGACTGAGGCTTCCAGCTGTGTACACCGGCTGGGTGCGCTGGCGGTGGGC 474
 Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTyrPheGluCys 211
 Db 475 AATGTGACACCGACATCTCCAGCAGGTGCTTTGGCCCTGGTCACTGTGAATCAGTGC 534

* * *

Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLys-----ThrPhe 227
 Db 535 -----CGCAGTACTGGGCTCAAGTATCACTGACTCCCATG 570
 Qy 228 LeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247
 Db 571 ATCTGTGCAGGT-----GGCGCAGGTGCTCTCGTCCAGGTGACTCCGGAGGCCCT 624
 Qy 248 LeuMetCysArgAsnLysLysGlyAlaTyrThrLeuAlaGlyValThrSerTyrPheLeu 267
 Db 625 CTCTGTCTGCCAG---AAGGGAACACATGGGTCTTATGTATGTCTCTCTGGGGCACC 681
 Qy 268 ---GlyCysGlyArgGlyTyrPheAsnValArgLysSerAspGlnGlySerProGly 286
 Db 682 AAAAATGCTC-----AATGTGCGC-----GCACCTGCT 708
 Qy 287 IlePheThrAspIleSerLysValLeuSerTyrPheLeuHis 299
 Db 709 GTGTATCTCGATTGAGTAGCAAGTTTCAGCACCTGGATCAAC 747
 RESULT 15
 LOCUS AK007765 876 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
 PRECURSOR (EC 3.4.21.1), full insert sequence.
 ACCESSION AK007765 GI:12841519
 VERSION AK007765.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to
 mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
 clone:1810044A17.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
 Yoneda,Y., Ishikawa,T., Tanaka,T., Tanaka,K., Matsuura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipillar sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
 Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
 Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,I.,
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
 Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
 Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
 Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,
 Quackenbush,J., Schriml,L.M., Staibli,F., Suzuki,R., Tomita,M.,
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
 Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,

Db 763 GTCACCTGCCCTCATGCCCCCTGGTTTCAGGAGATCTTGGAA 801

Search completed: March 25, 2003, 02:03:08
Job time : 1814 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 25, 2003, 02:03:14 ; Search time 104 Seconds
(without alignments)
2281.689 Million cell updates/sec

Title: US-09-735-713A-2
Perfect score: 1636
Sequence: 1 MSLKMLISRNKLLILLGIVF.....IFTDISKVLSTWIIHQTCN 306

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 538826 seqs, 387737923 residues
Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frames_p2n.model -DEV=xlh
-Q=/cgn2/uspto.spool/US09735713/runat_18032003_124143_21663/app_query.fasta_1.455
-DB=PublishedApplications_NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTEMP=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09735713 @CGN.1.1.57 @runat_18032003_124143_21663
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:
1: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2.6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	99.4	921	9	US-09-735-713A-1
2	1626	99.4	1568	9	US-09-735-713A-7
3	1608	98.3	909	9	US-09-735-713A-3
4	1372.5	83.9	1671	10	US-09-888-615-47

5	898	54.9	495	9	US-09-735-713A-5	Sequence 5, Appli
6	490.5	30.0	1327	9	US-09-978-295A-170	Sequence 170, App
7	490.5	30.0	1327	9	US-09-978-697-170	Sequence 170, App
8	490.5	30.0	1327	9	US-09-978-192A-170	Sequence 170, App
9	490.5	30.0	1327	9	US-09-999-832A-170	Sequence 170, App
10	490.5	30.0	1327	9	US-09-978-189-170	Sequence 170, App
11	490.5	30.0	1327	9	US-09-978-608A-170	Sequence 170, App
12	490.5	30.0	1327	9	US-09-978-191A-170	Sequence 170, App
13	490.5	30.0	1327	9	US-09-978-403A-170	Sequence 170, App
14	490.5	30.0	1327	9	US-09-978-564A-170	Sequence 170, App
15	490.5	30.0	1327	9	US-09-978-585A-170	Sequence 170, App
16	490.5	30.0	1327	9	US-10-017-081A-170	Sequence 170, App
17	490.5	30.0	3143	9	US-09-978-295A-168	Sequence 168, App
18	490.5	30.0	3143	9	US-09-978-697-168	Sequence 168, App
19	490.5	30.0	3143	9	US-09-978-192A-168	Sequence 168, App
20	490.5	30.0	3143	9	US-09-999-832A-168	Sequence 168, App
21	490.5	30.0	3143	9	US-09-978-189-168	Sequence 168, App
22	490.5	30.0	3143	9	US-09-978-608A-168	Sequence 168, App
23	490.5	30.0	3143	9	US-09-978-191A-168	Sequence 168, App
24	490.5	30.0	3143	9	US-09-978-403A-168	Sequence 168, App
25	490.5	30.0	3143	9	US-09-978-564A-168	Sequence 168, App
26	490.5	30.0	3143	9	US-09-978-585A-168	Sequence 168, App
27	490.5	30.0	3143	9	US-10-017-081A-168	Sequence 168, App
28	489.5	29.9	2409	10	US-09-888-615-54	Sequence 54, Appl
29	467	28.5	2412	9	US-09-978-295A-68	Sequence 68, Appl
30	467	28.5	2412	9	US-09-978-697-68	Sequence 68, Appl
31	467	28.5	2412	9	US-09-978-192A-68	Sequence 68, Appl
32	467	28.5	2412	9	US-09-999-832A-68	Sequence 68, Appl
33	467	28.5	2412	9	US-09-978-189-68	Sequence 68, Appl
34	467	28.5	2412	9	US-10-174-590-63	Sequence 63, Appl
35	467	28.5	2412	9	US-10-176-758-63	Sequence 63, Appl
36	467	28.5	2412	9	US-10-175-737-63	Sequence 63, Appl
37	467	28.5	2412	9	US-10-173-706-63	Sequence 63, Appl
38	467	28.5	2412	9	US-10-175-738-63	Sequence 63, Appl
39	467	28.5	2412	9	US-10-175-752-63	Sequence 63, Appl
40	467	28.5	2412	9	US-10-176-482-63	Sequence 63, Appl
41	467	28.5	2412	9	US-10-176-757-63	Sequence 63, Appl
42	467	28.5	2412	9	US-10-176-913-63	Sequence 63, Appl
43	467	28.5	2412	9	US-10-180-552-63	Sequence 63, Appl
44	467	28.5	2412	9	US-10-180-557-63	Sequence 63, Appl
45	467	28.5	2412	9	US-10-173-700-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-735-713A-1
; Sequence 1, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1e1 Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-1

Alignment Scores: 1.89e-170 Length: 921
Pred. No.: 1.89e-170

Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 9 Gaps: 0

US-09-735-713A-2 (1-306) x US-09-735-713A-1 (1-921)

```
Qy 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhe 20
Db 1 ATGAGTCTCAAAATGCTTATAGCAGGAAACAGCTGATTTACTACTAGGAATAGCTCTTT 60
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 61 TTGGAACRAGGTAAATCTGCARCTCTTTCGCTCCCAAGCTCCAGTGTGGGCAGAGT 120
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
Db 121 CTGGTTAAGGTACAGCCTTGGAAATTAATTTAACATTTTCAGTCGCATTTCTTGGAGGAAGC 180
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 181 CAAGTGGAGAGGGTCTCTACCTCGCAGGTATCTCTGAACAAAGCGAGAAGCATATT 240
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 241 TGTGAGGAGAGCATGCTCTACCACAGTGGGTGATCATCGGGGCTCAGTCGATTGCAAC 300
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Db 301 AGAAACATTTGTCTACTTTGAATCTTACTGCTGGAGAGTATGACTTAAGCCAGACACAC 360
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 361 CAGGAGAGCAAACTCTCACATTTGAAACTGTCTATCATCATCATCTTCTCCACCAAG 420
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 421 AAACCAATGGACTATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 480
Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 481 TTTGTGGGGCCCATATGTCCTCCAGAGCTGGGGAGCAATTTTGAAGCTGTGTTATTGT 540
Qy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValIleuGlnGlu 200
Db 541 ACAACTGCAGGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTGAGGAA 600
Qy 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 601 GTGAATCTGCCTATTTTTCACCTGGGAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
Qy 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 661 CCCATCAGTGGGAAGACCTTTCTTGCACAGGTTTCTGATGGAGGGAGAGACGCATGT 720
Qy 241 GlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Db 721 CAGGAGATTCAGGAGGTCTACTCATGTCCCGGAATAAGAAGGGGCTTGACTCTGGCT 780
Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
Db 781 GGTGTGACTTCTGGGGTTTGGGCTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAGT 840
Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Db 841 GATCAAGGATCCCTGGGATCTTACAGACATTAAGTAAAGTGTCTTCTGGATCCACGAA 900
Qy 301 HisIleGlnThrGlyAsn 306
Db 901 CACATCCAAACTGGTAAC 918
```

RESULT 2

US-09-735-713A-7
; Sequence 7, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US/09/735,713A
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-7

Alignment Scores:
Pred. No.: 4,14e-170 Length: 1568
Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 9 Gaps: 0

US-09-735-713A-2 (1-306) x US-09-735-713A-7 (1-1568)

```
Qy 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhe 20
Db 245 ATGAGTCTCAAAATGCTTATAGCAGGAAACAGCTGATTTACTACTAGGAATAGCTCTTT 304
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 305 TTTGAACRAGGTAAATCTGCARCTCTTTCGCTCCCAAGCTCCAGTGTGGGCAGAGT 364
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleuGlyGlySer 60
Db 365 CTGGTTAAGGTACAGCCTTGGAAATTAATTTAACATTTTCAGTCGCATTTCTTGGAGGAAGC 424
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 425 CAAGTGGAGGAAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAAGGCAGAGCATATT 484
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 485 TGTGAGGAGAGCATGCTCTACCACAGTGGGTGATCACGGGGCTCAGTCGATTGCAAC 544
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Db 545 AGAAACATTTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGC 604
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 605 CCAGGAGAGCAAACTCTCACATTTGAAACTGTCTATCATCATCATCTTCTCCACCAAG 664
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 665 AAACCAATGGACTATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 724
Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 725 TTTGTGGGGCCCATATGTCCTCCAGAGCTGGGAGCAATTTGAGGCTGGTATTATTGT 784
Qy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
```



```
Db 785 ACAATCGAGCTGGGGCGCTTAACCTGAAGTGGCGCTCTCTCACAACTTTCGACGAA 844
Qy 201 ValAsnLeuProIleLeuThrTrpGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 845 GTGAATCGCCCTATTTTGACCTGGAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 904
Qy 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 905 CCCATCAGTGGGAAGACCTTCTTTGCACAGGTTTCTGTATGAGGAGAGACGCGATGT 964
Qy 241 GlnGlyaspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Db 965 CAGGAGATTTCAGAGGTTTCACTCATGTGCGGGAATAAGAAAGGGCGCTGACCTCGCT 1024
Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
Db 1025 GGTGTGACTTCTGGGTTTGGGCTGTGTGCGAGCTGGAGAAACAATGTGAGGAAGT 1084
Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Db 1085 GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTGCTTCTCTGGATCCAGAA 1144
Qy 301 HisIleGlnThrGlyAsn 306
Db 1145 CACATCCAAACTGGTAAC 1162
RESULT 3
US-09-735-713A-3
; Sequence 3, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735.713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-3
Alignment Scores:
Pred. No.: 1.8e-168 Length: 909
Score: 1608.00 Matches: 300
Percent Similarity: 99.34% Conservative: 0
Best Local Similarity: 99.34% Mismatches: 2
Query Match: 98.29% Indels: 0
DB: 9 Gaps: 0
US-09-735-713A-2 (1-306) x US-09-735-713A-3 (1-909)
Qy 5 MetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhePheGluArgGly 24
Db 1 ATGCTTATAACGACGAAACAAGCTGATTTTACTAGTAATAGTCTTTTGAACRAGGT 60
Qy 25 LysSerAlaLeuLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal 44
Db 61 AAATCTGCARCTCTTTCGCTCCCAAGCTCCCAAGTGTGGGCGAGAGTCTGTTAAGGTA 120
Qy 45 GlnProIlePheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLys 64
Db 121 CAGCCTTGAATTTATTAACATTTTCAGTCGCATTTCTTGGAGGAAGCAAGTGGAGAA 180
```

```
Qy 65 GlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySer 84
Db 181 GGTTCCTATCCCTGGCAGGTATCTCTGAACAAGGCAAGACATATTTTGGAGGAAGC 240
Qy 85 IleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal 104
Db 241 ATCGTCTCACACACAGTGGGTGATCACGCGGCTCACATGCAATGCAAAACAGAAACATTTGTG 300
Qy 105 SerThrLeuAsnValThrAlaGlyClyTrpAspLeuSerGlnThrAspProGlyClyGln 124
Db 301 TCTACTTGAATGTACTGTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAA 360
Qy 125 ThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysProMetAsp 144
Db 361 ACTCTCACTATTGAACGTGCTCATCATCATCCACATTTCTCCACCAAGAAACCAATGAGAC 420
Qy 145 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 164
Db 421 TATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTTGTGGGGCCC 480
Qy 165 IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 184
Db 481 ATATGTCTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTTATTTGTACAACTGCAGGC 540
Qy 185 TrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnValAsnLeuPro 204
Db 541 TGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAACTTTCGAGGAAGTCAATCTGCCT 600
Qy 205 IleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIleSerGly 224
Db 601 ATTTTGACCTGGGAAGAGTGTGTGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGG 660
Qy 225 LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSer 244
Db 661 AAGACCTTCTTTCGACAGGTTTTCCTGATGGGAGAGACGCGATGTGAGGGAGATTCA 720
Qy 245 GlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSer 264
Db 721 GGAGTTCACATCATGTGCCGAATAAGAAAGGGCGCTGGACTCTGGCTGGTGTGACTTCC 780
Qy 265 TrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySer 284
Db 781 TGGGTTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCC 840
Qy 285 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 304
Db 841 CCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCTGGATCCACGAAACATCCAAACT 900
Qy 305 GlyAsn 306
Db 901 GGTAAC 906
RESULT 4
US-09-888-615-47
; Sequence 47, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOMMAN, GREGORY
; APPLICANT: WHITE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888.615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
```

```
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-47
```

Alignment Scores:

Pred. No.:	4.71e-142	Length:	1671
Score:	1372.50	Matches:	266
Percent Similarity:	93.75%	Conservative:	4
Best Local Similarity:	92.36%	Mismatches:	6
Query Match:	93.89%	Indels:	12
DB:	10	Gaps:	2

US-09-735-713A-2 (1-306) x US-09-888-615-47 (1-1671)

Qy	1	MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhe	20
Db	1	ATGAGTCTCAAAATGCTTATPAAGCAGCAACAAGCTGATTTTACTACTAGGAATAGTCTTT	60
Qy	21	PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer	40
Db	61	TTTGAACGAGTAAATCTGCACCTCTTTCCTCCCAAGCTCCCAGTTGTGGGCAGAGT	120
Qy	41	LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer	60
Db	121	CTGCTTAAAGTACAGCCTTGGAAATATTATTAACATTTTCAGTCGCATTTCTTGSAGGAAGC	180
Qy	61	GlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIle	80
Db	181	CAAGTGGAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAACAAAGGCAGACATATT	240
Qy	81	CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn	100
Db	241	TGTGGAGGAAGCATCTCTCACACAGTGGGTGATCACGGCGGCTCAGTCGATTCGAAC	300
Qy	101	ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAsp	120
Db	301	AGAAACATTTGTCTCTACTTTGAAATGTTACTGCTGGAGAGTATGACATTAAGCCAGACAGAC	360
Qy	121	ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys	140
Db	361	CCAGGAGAGCAAACTCTCACTATTGAACCTGTCTCATCATATCCACATTTCTCCACCAG	420
Qy	141	LysProMetAspTrpAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis	160
Db	421	AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC	480
Qy	161	PheValGlyProIleCysLeuProGluLeuArgGluClnPheGluAlaGlyPheIleCys	180
Db	481	TTTGTGGGGCCCATATGCTCTCCAGAGCTGCGGAGCAATTTGAGCCTGGTTTTATTGT	540
Qy	181	ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu	200
Db	541	ACAACTGCAGGCTGGGGCCGCTTAACCTGAAGGTGGCGCTCTCTCACAGCTCTTCGAGGAA	600
Qy	201	ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg	220
Db	601	GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG	660
Qy	221	ProIleSerGlyLysThrPheLeuCysThrGlyPheProAsnGlyClyArgAspAlaCys	240
Db	661	CCCATCAGTGGGAAGACCTTCTCTTTCACAGGTTTTCTTGATGGAGGGAGACGCATGT	720
Qy	241	GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaIleTrp-ThrLeuAl	260
Db	721	CAGGGAGATTACGAGGTTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGGACTCTGCG	780
Qy	260	aGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer	280
Db	781	TGCT-----CAATTTGG-----AGCTCAGGTGGGA807	

Qy 280 rAspGlnGlySerProGlyIle 287
|||::: | |||||:::|

Db 808 GGATCGCTTGAGTCCAGGAGTT 829

RESULT 5

```

US-09-735-713A-5
; Sequence 5, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-5

```

Alignment Scores:

Pred. No.:	2.13e-90	Length:	495
Score:	898.00	Matches:	164
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	54.89%	Indels:	0
DB:	9	Gaps:	0

US-09-735-713A-2 (1-306) x US-09-735-713A-5 (1-495)

QY	143	MetaSpTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal	162
DB	1	ATGCACTATGATATTGCCCTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTGTG	60
QY	163	GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr	182
DB	61	GGCCCCATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTTATTGTACAACCT	120
QY	183	AlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGluValAsn	202
DB	121	GCAGGCTGGGGCGCTTAACCTGAAGGTGGCGTCTCTCAACAGCTCTGCAGGAAGTGAAT	180
QY	203	LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle	222
DB	181	CTGCCATTTTGGACCTGGGAAGAGTGTGGCAGCTCTGTTAACACTAAAGAGGCCCATC	240
QY	223	SerGlyLysThrPheLeuCysThrGlyPheProAspGlyClyArgAspAlaCysGlnGly	242
DB	241	AGTGGGAAGACCTTCTTCTTCACAGGTTTTTCCTGATGGAGGGAGACGCATCTCAGGGA	300
QY	243	AspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyVal	262
DB	301	GATTCAGGAGTTCACTCATCTGCCGGATAAAGAAGGGGCCCTGGACTCTGGCTGGTG	360
QY	263	ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGln	282
DB	361	ACTTCTCGGGGTTTGGCGTGTGGCTGAGGCTGGAGAAACAATGTGAGGAAAAGTGATCAA	420
QY	283	GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle	302
DB	421	GGATPCCCTGGGATCTTCACAGACATTAGTAAAGTGCTTCTCTGGGATCCACGAACATC	480
QY	303	GlnThrGlyAsn	306
DB	481	CAAACTGGTAACT	492

RESULT 6
US-09-978-295A-170
; Sequence 170, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29

54 SerArgIleuGlyGlySerGlnValGluLysGlySerTyrProIleTrpGlnValSerLeu 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 534 AGCCGCATTGTTGGTGAGCTGTCTCCAGGGGTAGTGGCATGGCAGGCCAGCCTC 593
74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
QY

APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640

```
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.:          9,928-45      Length:      1327
Score:             490.50      Matches:    100
Percent Similarity: 56.42%      Conservative: 45
Best Local Similarity: 38.91%   Mismatches: 89
Query Match:       29.98%      Indels:     23
DB:                9          Gaps:         5

US-09-735-713A-2 (1-306) x US-09-978-697-170 (1-1327)

QY 54 SerAtgIleLeuGlyCysGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 534 ACCCGCATTTGTTGGAGCATCTGTCTCCGAGGGTGTAGTGGCCATGGCAGCCAGCCCTC 593

QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 594 CAGGTTCCGGGGTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATACAA 653

QY 94 AlaAlaHisCysIleAlaAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111
Db 654 GCTGCCACTGCTCCAGGAGGACAGCATGGCTCCACGGTGTCTGGACCGTGTTCCTG 713

QY 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 714 GCAAGGTGTGCGCAACCTCGCGCTGGCTGGAGAGGTCTCTCAAGGTGAGCGCGCTG 773

QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
Db 774 CTCCTGCAACCG--TACCACGAGAGGACACCCATGACTACGAGCTGGCGCTGTGCGAG 830

QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 831 CTCGACCAACCGGTGTGGCTCGCGCGCGCTGGCCCGCTGCTGCTGCCCGCGCGCTCC 890

QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 891 CACTTCTTGAGCCCGCGCTGCACACTGCTGGATACGGCTGGGGCGCTTGGCGAGGCG 950

QY 192 GlyValLeuSerGlnValLeuGlnValAsnValLeuProIleLeuThrTrpGluGluCys 211
Db 951 GSCCCCATCAGCAGCGCTCTCGAAGATGGATGTGCAGTTGATCCCAAGGACCTGTGC 1010

QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 1011 AGCGAGGCTATCTCGTACCAGGTGAGCCCA-----CGCATGCTGTGTGCCGCG 1058

QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
Db 1059 TACCGCAAGGCAAGAGGATGCTGTACGGGTGACTCAGGTGGTCCGCTGGTGTGCAAG 1118

QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACCTCAGTGGCGCTGGTTCCTGGCGGGCGTGGTCAGCTGGGCGCTGGCTGGCCCG 1178

QY 272 GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTCGGCGCTAC 1199

QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCACAGGTGTGATCATGCTGGATCCAGCAAGTGGTGGTACCTGAGGA 1250

RESULT 8
US-09-978-192A-170
; Sequence 170, Application US/09978192A
; Patent No. US2002017755A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: ROY, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
```


; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:	9.92e-45	Length:	1327
Score:	490.50	Matches:	100
Percent Similarity:	56.42%	Conservative:	45
Best Local Similarity:	38.91%	Mismatches:	89
Query Match:	29.98%	Indels:	23
DB:	9	Gaps:	5

US-09-735-713A-2 (1-306) x US-09-978-192A-170 (1-1327)

Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 534 AGCCGCAATGTTGGAGCTGTCTCCGAGGGTGAGTGGCCATGCGAGCCAGCCCTC 593
Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 594 CAGGTTCCGGGTCGACACATCTGTGGGGGCGCCCTCATCGCTGACCGCTGGGTGATAACA 653
Qy 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111
Db 654 GCTGCCCACTGCTCCAGAGAGACAGCATGGCCCTCCAGCGGTGCTGTGGACCGTGTCTCG 713
Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 714 GCGAAGGTGGCAGAACTCGCGTGGCTGGAGAGGTGTCTTCAAGGTGAGCGCGCTG 773
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
Db 774 CTCCTGCACCCG--TACCAGGAAGAGACACCATGACTACGACGTGCGCTGCTGCAG 830
Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 831 CTCGACCCCGGTGGTGGCTGGCCCGCGCTGGCCCGCTGCTGCTGCGCGCGCTCC 890
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 891 CACTTCTCGAGCCGGCTGCATCTGTGGATTAGGGCTGGGGCGCTTCCGGGAGGCG 950
Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
Db 951 GGCCCATCAGCAAGCTCTCAGAAAGTGTGATGTCAGTTGATCCACAGGACCTGTGC 1010
Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 1011 AGCGAGGCTTATCGCTACAGGTGACGCCA-----CGCATGCTGTGTGCCCGC 1058
Qy 232 PheProAspGlyGlyArgAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
Db 1059 TACCGAAGGCAAGAGATGCTGTGAGGTGACTCAGTGTGCTGCTGCTGTGCAAG 1118
Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACCTAGTGGCGCTGTCTTCCGCGGGCTGTGTCAGCTGGGCGCTGGCTGTGGCCG 1178
Qy 272 GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTCGGCGCTTAC 1199
Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCAGGTGTGATCAGTGGATCCAGCAAGTGGTACCTGAGGA 1250

RESULT 9

US-09-999-832A-170
; Sequence 170, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David
; APPLICANT: Desnayers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PLC63
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786

; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 9.92e-45 Length: 1327
Score: 490.50 Matches: 100
Percent Similarity: 56.42% Conservative: 45
Best Local Similarity: 38.91% Mismatches: 89
Query Match: 29.98% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-2 (1-306) x US-09-999-832a-170 (1-1327)

QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTlpGlnValSerLeu 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 534 AGCGCATTTGTTGGTGAGCTGTCTCCGAGGGTGAGTGGCCATGGCAGCCGAGCTC 593
QY 74 LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTlpValIleThr 93
::: ||| ::|||:|||||:|||||:|||||:|||||:|||||:|||||
Db 594 CAGGTTCCGGGTGCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATACA 653
QY 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 654 GCTGCCCACTGCTTCCAGGAGGACAGCATGGCTCCACGGTGTGTGGACCGTGTCTCTG 713
QY 112 GlycIuTyAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131

```
Db 714 GGCAAGGTGGCAGAACTCGCCTGGCTGGAGAGGTGCTTCAAGGTAGCCGCCGTG 773
QY 132 IleIleHisProHisSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
Db 774 CTCCTGCACCCG--TACCAGGAAGAGCAGACCATGACTACGACGTGGCGTGTGCAG 830
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 831 CTCGACACCCGGTGTGGCTCGCGCGCGCTGCGCGCGCTGCTGCGCGCGCTCC 890
QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgIleThrGluGly 191
Db 891 CACTTCTTGAGCCCGCCTGCACTGCTGATACGGGTGGGCGCCCTGGCGGAGGCG 950
QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
Db 951 GGCCCATCAGCAACCTCTGCAGAAAGTGGATGTCAGTTGATCCACAGACCTGTGC 1010
QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 1011 AGCGAGGCTATCGCTACCAGGTGAGGCCA-----CGCATGCTGTGTGCGCGC 1058
QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyClySerLeuMetCysArg 251
Db 1059 TACCGCAAGGCAAGAAAGATGCTGTACGGGTGACTCAGGTGGTCCGCTGGTGTCAAG 1118
QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACCTAGTGGCGCTGTCTCTGGCGGGCTGGTCACTGAGTGGCGCTGGCTGTGGCGG 1178
QY 272 GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTCGGCGCTCTAC 1199
QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCAGGTGTATCATGCTGGATCCAGCAAGTGGTACCTGAGGA 1250

RESULT 10
US-09-978-189-170
; Sequence 170, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2630PIC7
; CURRENT APPLICATION NUMBER: US/09/978,189
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
```



```

Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACCTAGTGGCGCTGCTCTCCGCGGGCTGCTCAGCTGGGCGCTGGGCGG 1178
Qy 272 GlyTrpArgAsnAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTCGGCGCTCTAC 1199

Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTGACCTGAGGA 1250

RESULT 11
US-09-735-713a-2 (1-306) x US-09-735-713a-2 (1-306)
; Sequence 170, Application US/09978608A
; Publication No. US20030045462A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C22
; CURRENT APPLICATION NUMBER: US/09/978,608A
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 170
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713a-2 (1-306) x US-09-735-713a-2 (1-306)

Alignment Scores:
Pred. No.: 9,92e-45 Length: 1327
Score: 490.50 Matches: 100
Percent Similarity: 56.42% Conservative: 45
Best Local Similarity: 38.91% Mismatches: 89
Query Match: 29.98% Indels: 23
DB: 5 Gaps: 5

US-09-735-713a-2 (1-306) x US-09-735-713a-2 (1-306)

Qy 54 SerArgIleLeuGlySerGlnValGluLysGlySerTrpTrpGlnValSerLeu 73
Db 534 AGCGCATGTGTGGAGCTGTCTCCAGAGGTGAGTGCCATGCGCAGCGCCGCTC 593
Qy 74 LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTrpValIleThr 93

```

```

Db 594 CAGGTTGGGGTGCACACATCTGTGGGGGGCCCTCATCTGCTGACCGCTGGGTGATAACA 653
Qy 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111
Db 654 GCTCCCACTGCTTCCAGGAGGACAGCATGCCTCCAGCGTGTCTGGACCGCTGTCTCCTG 713
Qy 112 GlyGluTrpAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 714 GCGAGGTGTGGCAGACTCCGCTGCGAGAGGTGCTTCAAGGTGAGCGCGCTG 773
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys 151
Db 774 CTCCTGCACCG---TACCACGAAGAGGACACCATGATGACGAGCTGCGCTGCGCAG 830
Qy 152 MetaAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 831 CTCGACACCGCGGTGGTGGCTGCGCGCGCTGCGCGCGCTGCTGCTGCGCGCGCTCC 890
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 891 CACTTCCTCGAGCCGCGCTGCTGCTGGATTACGGGCTGGGGCGCTTGGCGAGGCG 950
Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
Db 951 GCGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTCAGTTGATCCACACAGACCTGTGC 1010
Qy 212 ValAlaAlaLeuLeuThrLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 1011 AGCGAGGCTATCTGCTACCGAGTGCAGCCA-----CGATGCTGTGCGCGCG 1058
Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
Db 1059 TACCGCAAGGCGCAAGAGGATGCTGTCAGGGTGACTCAGCTGCTGCTGCTGCAAG 1118
Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACCTAGTGGCGCTGCTCTCCGCGGGGCTGCTGCTGCGCGCTGGGCTGGGCGG 1178
Qy 272 GlyTrpArgAsnAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTCGGCGCTCTAC 1199

Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTGACCTGAGGA 1250

RESULT 12
US-09-735-713a-2 (1-306) x US-09-735-713a-2 (1-306)
; Sequence 170, Application US/09978608A
; Publication No. US20030045462A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

```

APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C4
CURRENT APPLICATION NUMBER: US/09/978,191A
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 9, 92e-45 Length: 1327
Score: 490.50 Matches: 100
Percent Similarity: 56.42% Conservatives: 45
Best Local Similarity: 38.91% Mismatches: 89
Query Match: 29.98% Indels: 23
DB: 9 Gaps: 5

US-09-735-713A-2 (1-306) x US-09-978-403A-170 (1-1327)

QY 54 SerArgIleLeuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 534 AGCGGCGATTGTTGGTGGAGCTGTCTCCGAGGGTGTAGTGGCCATGGCAGCCAGCCCTC 593
QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 594 CAGGTTCGGGGTTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGTGATAACA 653
QY 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111
Db 654 GTGCGCCACTGTTCGAGGAGACGATGGCTCCACGGTGTGGACCGTGTTCCTG 713
QY 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 714 GCGAAGGTGGCAGAACTCGCGCTGGCTGGAGAGGTGTCTTCAAGGTGAGCGCCCTG 773
QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys 151
Db 774 CTCCTGCACCCG---TACCACGAAGAGACAGCCATGATGACGAGTGGCGTGTGCAG 830
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 831 CTCGACCACCGGTGGCTGGCGCGCTGGCGCGCTGTGCTGCCCGCGCGCTCC 890
QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 891 CACTTCTTCGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluLys 211
Db 951 GCGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTCAGTTGATCCACAGGACCTGTGC 1010
QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 1011 ACGGAGGCTATCGTACCAGGTGAGCCA-----CGCATGCTGTGTGCCGCGC 1058
QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlySerLeuMetCysArg 251
Db 1059 TACCGCAAGGGCAAGAGGATCCCTGTGAGGTGACTCAGGTGGTCCGCTGGTGTGCAAG 1118
QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACACTGAGTGGCGCTGTCTCTGGGGGGTGGTGCAGCTGGGGCTGGCTGTGGCGGG 1178
QY 272 GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTGGCGCTGTAC 1199
QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCAGAGGTGTATGATGCTGGATCCAGCAAGGTGAGTACCTGAGGA 1250

RESULT 14

US-09-978-564A-170
; Sequence 170, Application US/09978564A
; Publication No. US20030050241A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David

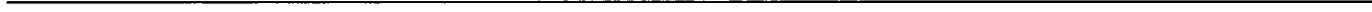
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: ROY, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC25
; CURRENT APPLICATION NUMBER: US/09/978,564A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27

Db	714	GGCAAGGTGTGGCAGAACTCGCGCTGGCTGGAGAGGTGTCTCTCAAGGTGAGCGCGCTG	773
Qy	132	IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys	151
Db	774	CTCTGCAACCG--TACCACGAAGAGGACACCCATGACTACGACGTGGCGCTGCTGCAG	830
Qy	152	MetaAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171
Db	831	CTCGAACCAACCGGTGTGGCTCGGCGCGCGTGGCCCGCTGCTGCTGCCCGCGCGCTCC	890
Qy	172	GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	191
Db	891	CACTTCTTCAGCGCGCGCTGCACTGCTGGATTACGGGCTGGGCGCCTTCGCGGAGGC	950
Qy	192	GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys	211
Db	951	GGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGGACCTGTGC	1010
Qy	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231
Db	1011	AGCGAGGCGCTATCGCTACCAGGTGAGGCCA-----CGCATCTGTGTGCCGCG	1058
Qy	232	PheProAspGlyClyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251
Db	1059	TACCGCAAGGGCAAGAGGATGCTCTCAGGTGACTCAGGTGGTCCGCTGGTGTGCAAG	1118
Qy	252	AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg	271
Db	1119	GCACTCAGTGGCGCGTGTCTCGCGGCGCTGGTCAGCTGGGCGCTGGGCTGGGCGGG	1178
Qy	272	GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe	288
Db	1179	-----CCTAACTACTCGGCGCTAC	1199
Qy	289	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly	305
Db	1200	ACCCGATCAAGGTGTGATCAGCTGGATTCAGCAAGTGGTGCACCTCAGGA1250	

RESULT 15

US-09-978-585A-170
 ; Sequence 170, Application US/09978585A
 ; Publication No. US20030049633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillen, Kenneth J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2630PIC15

Search completed: March 25, 2003, 03:56:34
Job time : 117 secs



THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 25, 2003, 00:14:23 ; Search time 271 Seconds
(without alignments)
2542.847 Million cell updates/sec

Title: US-09-735-713A-2
Perfect score: 1636
Sequence: 1 MSLSKMLISRNKLLILLGIVE.....IFTDISKVLSWIHIQTGN 306

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US0735713/runat_18032003_124141_21580/app_query.fasta_1.455
-DB=N.Geneseq_101002 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0735713 @CGN_1.1_263 @runat_18032003_124141_21580 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1632	99.8	1262	24	ABK12903	Human protease PRT
2	1626	99.4	921	22	AA09328	Human protease hom
3	1626	99.4	1568	22	AA09331	Human protease hom
4	1608	98.3	909	22	AA09329	Human protease hom
5	1372.5	83.9	1671	24	ABK31790	DNA encoding novel
6	950	58.1	637	22	AAH98934	Toad EST-derived c
7	898	54.9	495	22	AA09330	Human protease hom
8	567.5	34.7	2742	23	AA588433	DNA encoding novel
9	567.5	34.7	3309	23	AA587759	DNA encoding novel
10	567.5	34.7	4729	23	AA572057	DNA encoding novel
11	567.5	34.7	4729	23	AA574405	DNA encoding novel
12	567.5	34.7	4729	23	AA578775	DNA encoding novel
13	567.5	34.7	4729	23	AA584074	DNA encoding novel
14	561.5	34.3	3096	24	ABQ86158	Novel human gene.
15	490.5	30.0	1327	20	AA234034	Human EST DNA35597
16	490.5	30.0	1327	21	AACT78455	Human EST DNA35597
17	490.5	30.0	1327	21	AA58237	Human EST (express
18	490.5	30.0	3143	20	AA234033	Human PRO618 nucle
19	490.5	30.0	3143	21	AACT78494	Human PRO618 (UNQ3
20	490.5	30.0	3143	21	AA58236	Human PRO618 nucle
21	489.5	29.9	2409	24	ABK31797	DNA encoding novel
22	486.5	29.7	2672	22	AA013117	Human membrane-typ
23	486.5	29.7	3104	22	AA013116	Human membrane-typ
24	486	29.7	3736	23	ABL28971	Drosophila melanog
25	476.5	29.1	708	22	AA013115	Human membrane-typ
26	467	28.5	1656	22	AA013118	Human membrane-typ
27	467	28.5	2412	20	AA233949	Human PRO382 nucle
28	467	28.5	2412	21	AACT78475	Human PRO382 (UNQ3
29	467	28.5	2412	22	AA545956	Human DNA encoding
30	467	28.5	2948	24	AA036726	Human trypsin fami
31	466	28.5	2267	24	ABN95751	Gene #2249 used to
32	466	28.5	2267	24	ABL67214	Thyroid cancer rel
33	466	28.5	2434	23	ABV22421	Human prostate exp
34	466	28.5	2434	23	ABV28235	Human prostate exp
35	462	28.2	2413	21	AA593842	Tumour associated
36	462	28.2	2413	22	AA599907	Human cDNA encodin
37	456.5	27.9	1365	24	ABK31787	DNA encoding novel
38	456.5	27.9	1394	21	AA02323	Human serine prote
39	456.5	27.9	1394	24	AA028649	Human serine prote
40	456.5	27.9	2123	24	ABQ55071	Human ovarian anti
41	456.5	27.9	2951	24	AA036727	Human trypsin fami
42	454.5	27.8	1697	21	AACT7781	Human cancer assoc
43	454.5	27.8	3387	24	ABK31781	DNA encoding novel
44	454.5	27.8	3711	24	ABK86141	cDNA encoding huma
45	448	27.4	927	24	ABK16740	Epithelin-like ser

ALIGNMENTS

RESULT 1

ABK12903
ID ABK12903 standard; cDNA; 1262 BP.

AC ABK12903;

DT 09-APR-2002 (first entry)

DE Human protease PRTS-20 cDNA sequence.

XX Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
XX cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX Homo sapiens.
OS

XX Key Location/Qualifiers
 FH 165..1085
 FT /*tag= a
 FT /product= "Human protease PRTS-20"
 XX
 PN WO200198468-A2.
 XX
 XX 27-DEC-2001.
 XX
 XX 13-JUN-2001; 2001WO-US19178.
 XX
 XX 16-JUN-2000; 2000US-212336P.
 PR 22-JUN-2000; 2000US-213955P.
 PR 29-JUN-2000; 2000US-215396P.
 PR 07-JUL-2000; 2000US-216821P.
 PR 14-JUL-2000; 2000US-218946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walla NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azinzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;
 XX
 XX WPI; 2002-090437/12.
 XX P-PSDB; AAU74760.
 XX
 XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
 PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
 PT proliferative (e.g. cancer) disorders -
 XX
 XX Claim 5; Page 175; 177pp; English.
 XX
 XX The present invention relates to twenty one new human proteases,
 CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
 CC polypeptides of the invention are useful in the diagnosis, treatment and
 CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
 CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
 CC myocardial infarction, autoimmune/inflammatory e.g. acquired
 CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
 CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
 CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
 CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
 CC endometriosis disorders. Numerous other examples of each disorder are
 CC given in the specification. The present nucleic acid sequence encodes
 CC the human protease PRTS-20 protein of the invention.
 XX
 XX SQ Sequence 1262 BP; 354 A; 265 C; 306 G; 337 T; 0 other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 6,74e-149 Length: 1262
 XX Score: 1632.00 Matches: 305
 XX Percent Similarity: 99.67% Conservative: 0
 XX Best Local Similarity: 99.67% Mismatches: 1
 XX Query Match: 99.76% Indels: 0
 XX DB: 24 Gaps: 0
 XX
 XX US-09-735-713A-2 (1-306) x ABRK12903 (1-1262)
 XX
 XX QY 1 MetSerLeuLysMetLeuSerArgAsnLysLeuLeuLeuLeuGlyIleValPhe 20
 Db 165 ATGAGTCTCAAAATGCTTAAGCAGGAAACAAGCTGATTTTACTACTAGTAATAGTCTTT 224
 QY 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
 Db 225 TTGTACAGGAGGAAATCTCCAACTCTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 284
 QY 41 LeuValLysValGlnProTyrPheAsnLysPheAsnLysPheSerArgIleLeuGlyLysSer 60
 Db 285 CTGGTTAAGTACAGCTTGGATTAATTTTAACTATTTTACATTTTCAGTCGCAATCTTGGAGGAGC 344

QY 61 GlnValGluLysGlySerTyrProTyrGlnValSerLeuLysGlnArgGlnLysHisIle 80
 Db 345 CAAAGTGGAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGAGCAAGCATATT 404
 QY 81 CysGlyGlySerIleValSerProGlnTyrValIleThrAlaAlaHisCysIleAlaAsn 100
 Db 405 TGTGGAGGAAGCATCGTCTCACACAGTGGTGTATCAGCGCGCTCAGTCGATTGCAAC 464
 QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
 Db 465 AGAAACATTGTCTACTTGAATGTACTGTGGAGATGATGACTTAAGCCAGACAGAC 524
 QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
 Db 525 CCAGGAGAGCAAACTCTCACTATTGAAACTGTATCATCATCATCATTTCTCCACCAAG 584
 QY 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
 Db 585 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 644
 QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
 Db 645 TTGTGGGGCCCATATGTTCTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGT 704
 QY 181 ThrThrAlaGlyTyrPglyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
 Db 705 ACAACTGCAGGCTGGGGCGGCTTAAGTGAAGTGGGCTCTCTCACAAGTCTTCAGGAA 764
 QY 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
 Db 765 GTGAATCTGCCCTATTTTGACCTGGGAAGAGTGTGGCAGCTCTGTTAAACACTAAAGAGG 824
 QY 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
 Db 825 CCATCAGTGGGAAGACCTTTCTTGACAGGTTTCTCGTGGAGGGAGACCGCATGT 884
 QY 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaThrThrLeuAla 260
 Db 885 CAGGAGAGTTCAGGAGGTTCACTCATGTCCCGAATAAGAAAGGGCGCTGGACTCTGGCT 944
 QY 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
 Db 945 GGTGTGACTTCTCGGGGTTGGGCTGTGTCGAGGCTGGAGAACAATGTGAGGAAAGT 1004
 QY 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
 Db 1005 GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTCTTTCTCTGGATCCACGAA 1064
 QY 301 HisIleGlnThrGlyAsn 306
 Db 1065 CACATCCAAACTGGTAAAC 1082
 XX
 XX RESULT 2
 XX AAD09328
 XX ID AAD09328 standard; cDNA; 921 BP.
 XX AC AAD09328;
 XX AC AAD09328;
 XX DT 10-SEP-2001 (first entry)
 XX
 XX Human protease homologue cDNA #1.
 XX
 XX Human; protease homologue; novel human protein; NHP; therapy;
 KW pharmacogenomic application; physiological disorder; ss.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..921
 FT CDS /*tag= a
 FT /product= "Human protease homologue #1"
 FT unsure 67..69

[illegible]

XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-408641/43.
XX
XX Polynucleotide encoding novel human protease homologs, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
XX Disclosure; Page 31; 32pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein (NHP),
CC known as human protease homologue. NHP shares structural similarity
CC with animal proteases, particularly trypsin-like protease such
CC as oviductin, plasminogen activator and human plasma kallikrein
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
CC pharmacogenomic applications. NHP sequences are useful for identifying
CC agonists, antagonists and modulators and also for producing antibodies
CC useful in diagnosis, drug screening, clinical trial monitoring and in
CC treatment of physiological disorders.
XX
SQ Sequence 1568 BP; 452 A; 338 C; 353 G; 423 T; 2 other;

Alignment Scores:
Pred. No.: 3,43e-148 Length: 1568
Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 22 Gaps: 0

US-09-735-713A-2 (1-306) x AAD09331 (1-1568)

QY 1 MetSerLeuLysMetLeuLLeSerArgAsnLysLeuLLeLeuLeuLysValPhe 20
DB 245 ATGAGTCTCAAAATGCTTATAGCAGAACAGCTGATTTTACTACGAGTAATGCTTT 304

QY 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
DB 305 TTTGAACRAGGTAATTCGACRCCTTTTCGCTCCCAAGCTCCAGTGTGGCAGAGT 364

QY 41 LeuValLysValGlnProTrpAsnTyPheAsnIlePheSerArgIleLeuGlyGlySer 60
DB 365 CTGTTTAAGGTACAGCCTTGAATATTTTAAACATTTTTCAGTCGCATCTTGGAGGAGC 424

QY 61 GlnValGluLysGlySerTyProTrpGlnValSerLeuLysGlnArgGlnLysHistle 80
DB 425 CAAGTGGAGAAGGGTTCTATCCCTGGCAGGTATCTCTGAACAAAGGCAGACATATT 484

QY 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
DB 485 TGTGGAGAGCATCGCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAC 544

QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyArgPheSerGlnThrAsp 120
DB 545 AGAAACATTTGTCTACTTTGAATGTTACTGCTGAGAGTATGACTTAAAGCCAGACAGAC 604

QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
DB 605 CCAGGAGAGCAACTCTCATATTGAACATGTCATCATCATATCCACATTTCTCCACCAAG 664

QY 141 LysProMetAspTyArgPheIleAlaLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
DB 665 AAACCATGGACTATGATATTTGCCCTTTGAGATGGCTGGAGCCTTCCAAATTTGGCCAC 724

QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
DB 725 TTTCTGGGGCCCATATCTCTCCAGAGCTGCGGAGCAATTTGAGGCTGTTTATTTGT 784

QY 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyClyValLeuSerGlnValLeuGlnGlu 200
DB 785 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGGTGGCGCTCTCTCACAAGTCTTTCAGGAA 844

QY 201 ValAsnLeuProIleLeuThrTrpGluCysValAlaAlaLeuLeuThrLeuLysArg 220
DB 845 GTGAATCTGCTATTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 904

QY 221 ProfileSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
DB 905 CCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTCTCTGATGGAGGAGAGAGCATGT 964

QY 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
DB 965 CAGGAGATTTCAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCCTGGACTTGGCT 1024

QY 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280
DB 1025 GGTCTGACTCTCTGGGTTTGGGCTGTGGTCAGGCTGGAGAACATGTGAGGAAAAGT 1084

QY 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
DB 1085 GATCAAGGATCCCTGGGATCTTCACAGACATATTAGTAAAGTGCTTCTCTGGATCCACGA 1144

QY 301 HistleGlnThrGlyAsn 306
DB 1145 CACATCCAAACTGGTAAC 1162

RESULT 4
AAD09329
ID AAD09329 standard; cDNA; 909 BP.
XX
AC AAD09329;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human protease homologue cDNA #2.
XX
KW Human; protease homologue; novel human protein; NHP; therapy;
KW pharmacogenomic application; physiological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..909
FT /*tag= a
FT /product= "Human protease homologue #2"
FT 55..57
FT /*tag= b
FT /note= "Encodes Arg"
FT misc_feature 56
FT /*tag= c
FT /note= "This degenerate base represents a
FT polymorphic site"
FT 70..72
FT /*tag= d
FT /note= "Encodes Ala"
FT 70
FT misc_feature 70
FT /*tag= e
FT /note= "This degenerate base represents a
FT polymorphic site"
XX
PN WO200146407-A1.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US33738.
XX
PR 23-DEC-1999; 99US-0171566.
XX (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2001-408641/43.
DR P-PSDB; AAE04734.
DR

CC hypertension, psychotic disorders, neurological disorders
CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
CC The nucleic acids and polypeptides are also useful for treating viral
CC infections caused by human immunodeficiency virus (HIV), and non-viral
CC infections such as ocular disease (e.g. glaucoma) and macular
CC degeneration. ABK31744-ABK31802 represent DNA sequences encoding for
CC the novel human proteases of the invention.
XX

SQ Sequence 1671 BP; 451 A; 369 C; 404 G; 447 T; 0 other;

Alignment Scores:

Pred. No.: 1.54e-123 Length: 1671
Score: 1372.50 Matches: 266
Percent Similarity: 93.75% Conservative: 4
Best Local Similarity: 92.36% Mismatches: 6
Query Match: 83.89% Indels: 12
DB: 24 Gaps: 2

US-09-735-713A-2 (1-306) x ABK31790 (1-1671)

QY 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhe 20
DB 1 ATGAGTCTCAAAATGCTTATAGCAGAACAGCTGATTTTACTACTAGGAATAGTCTTT 60
QY 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
DB 61 TTTGACGAGGTAATCTGCAACTCTTTCCGCTCCCAAGCTCCAGTTGTGGCAGAGT 120
QY 41 LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyLys 60
DB 121 CTGGTTAAGTACAGCTTGGAAATATTTTAACTTTTCAGTGGCATTCTTGGAGGAAGC 180
QY 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
DB 181 CAAAGTGGACAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGCGAGCATATT 240
QY 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
DB 241 TGTGGAGGAAGCATCTCTCACACAGTGGTGTATCAGCGCGCTCACATGCAATTCGAAAC 300
QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
DB 301 AGAAACATTGTCTACTTTGAATGTACTGTGGAGAGTATGACTTAAAGCCAGACAGAC 360
QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
DB 361 CCAGGAGACCAACTCTACTATTGAACTGTATCATCATCATCATCTCCACCAAG 420
QY 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
DB 421 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 480
QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
DB 481 TTTGTGGGCCCATATGTTCTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTGT 540
QY 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
DB 541 ACAACTGCAGGCTGGGCGCTTAAGTGAAGTGGCTCTCTCAAGCTCTTGAGGNA 600
QY 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
DB 601 GTGAATCTGCTTATTTTACCTTGGGAAGAGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
QY 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
DB 661 CCCATCAGTGGGAACCTTTCTTGCACAGGTTTCTGTATGGAGGAGACCATGT 720
QY 241 GlnGlyAspSerGlyLysSerLeuMetCysArgAsnLysLysGlyValAlaTrp-ThrLeuAl 260
DB 721 CAGGAGATTCAGAGGTTCACTCATGTGCCGGAATAGAAAGGGCCCTGGGACTCTGGC 780
QY 260 aglyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSe 280

DB 781 TGGT-----CAATTTGGG-----AGGCTCAGGTGGGA 807
QY 280 rAspGlnGlySerProGlyIle 287
DB 808 GGATCGCTTGGTCCAGTCCAGGAGTT 829
RESULT 6
AAH98934
ID AAH98934 standard; cDNA; 637 BP.
XX
AC AAH98934;
XX
DT 12-OCT-2001 (first entry)
DE Toad EST-derived coding sequence SEQ ID NO: 791.
XX

Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
OS Xenopus laevis.

PN WO200154477-A2.

PD 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02687.

XX 25-JAN-2000; 2000US-0491404.

PR 17-JUL-2000; 2000US-0617746.

PR 03-AUG-2000; 2000US-0631451.

PR 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;

PI Cao Y, Drmanac RA, Zhang J, Werhman T;

XX WPI; 2001-476164/51.

DR P-PSDB; AAM24275.

PT Isolated polypeptide for treatment of diseases, diagnostics, raising

PT antibodies and research use -

XX Claim 1; Page 673-674; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
proteins from a variety of organisms, including human, dog, cat, horse,
cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
urchin and tomato. These were derived from expressed sequence tags (ESTs)
from the organism of interest. They can be used in diagnostics,
forensics, gene mapping, identification of mutations, to assess
biodiversity and for nutritional purposes. The present sequence is a cDNA
of the invention.

SQ Sequence 637 BP; 172 A; 133 C; 141 G; 191 T; 0 other;

Alignment Scores:

Pred. No.: 4.62e-83 Length: 637
Score: 950.00 Matches: 181
Percent Similarity: 95.00% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 58.07% Indels: 0
DB: 22 Gaps: 0

US-09-735-713A-2 (1-306) x AAH98934 (1-637)

QY 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20

DB 91 ATGAGTCTCAAAATGCTTATAGCAGGAACAAGCTGATTTTACTAGGAATAGTCTTT 150

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX WO200175067-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG24246.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 1; SEQ ID NO 24237; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2742 BP; 835 A; 561 C; 664 G; 682 T; 0 other;

Alignment Scores:

Pred. No.: 4.35e-45 Length: 2742
 Score: 567.50 Matches: 114
 Percent Similarity: 59.57% Conservative: 51
 Best Local Similarity: 41.16% Mismatches: 101
 Query Match: 34.69% Indels: 11
 DB: 23 Gaps: 4

US-09-735-713A-2 (1-306) x AAS88433 (1-2742)

Qy 37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTyrPheAsnIlePhe 53
 Db 25 TGTGGAATTCGCATGGTCAACATCAAAAGTAAGGAACCTGCGGTGGGATCTAGATTCTC 84
 Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
 Db 85 TCTAGAATTAGTAGTTGGAGAAATTCACACAGTGTGACTGCATCCATGCGAGGTCTCCCTA 144
 Qy 74 LysGlnGlnGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
 Db 145 AATCAGATGAGCACCACCTCTGTGGAGGAGGCTTGATTCAAGAGATCGGGTGTGTACA 204

Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
 Db 205 GCAGCACACTGCCTGGACAGCCTCAGTCAGAGCAGCTGGAAGATATAAAGTGTGACTTCT 264
 Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
 Db 265 GGGGAGTACAGCCTCTTTTTCAGAGGATAAGCAAGAACAGATATATCTCTGTCTCAAAAT 324
 Qy 132 IleIleHisProHisPheSerThrLysProMetAspTyrAspIleAlaLeuLys 151
 Db 325 ATTACCATCTGTAATACAGCCGCTGAATATATAGTCTCTGATATTTGCACTGCTGTAT 384
 Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
 Db 385 CTAAACACAAAGTCAAGTTTGGAAATCTCTTCAGCCAACTGCTCTTCAGACGCGAT 444
 Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
 Db 445 GATAAAGTTGAACCAAGCAAGAAATCTTTTGTCTTATCCAGTGTGATGGGCAAGATTTCCAACA 504
 Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
 Db 505 TCAGAAATATTCAAATGCTCTACAAAGAAATGGAATCTCCATCATGGATGACAGACGCTGT 564
 Qy 212 ValAlaAlaLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
 Db 565 AATACTGTGCTCAAGACGATGAACCTCCCTCCCTGGGAAGACCATCTGTGTGCTGGC 624
 Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
 Db 625 TTCCCTGATTGGGAATGGAGCGCTGCCAGGGGACTCTGGAGGACCACTGGTTGTAGA 684
 Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
 Db 685 AGAGTGGTGGAAATCTGGATTTCTGGGGATAACTCTCTGGGTAGCTGGTGTGCTGGA 744
 Qy 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
 Db 745 GGTTCAGTTCCTCCGTGAAGAACACCATCTGAAGCA-----TCACCTTGGCATTTTC 795
 Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
 Db 796 TCCAAAGTGTCTGATGATGATTTATCATCTCAAAACCTGTTTCACAGGT 846
 RESULT 9
 AAS87759/C
 ID AAS87759 standard; cDNA; 3309 BP.
 XX AC AAS87759;
 XX DT 13-FEB-2002 (first entry)
 XX DE DNA encoding novel human diagnostic protein #23563.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX

DR WPI: 2001-639362/73.
 DR P-PSDB; ABG23572.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID NO 23563; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3309 BP; 789 A; 783 C; 683 G; 1054 T; 0 other;

Alignment Scores:

Pred. No.:	5,56e-45	Length:	3309
Score:	567.50	Matches:	114
Percent Similarity:	59.57%	Conservative:	51
Best Local Similarity:	41.16%	Mismatches:	101
Query Match:	34.69%	Indels:	11
DB:	23	Gaps:	4

US-09-735-713A-2 (1-306) x AAS87759 (1-3309)

Qy	37	CysGlyGlnSerLeuVal-----GlnProTrpAsnTyrPheAsnIlePhe	53
Db	3285	TGTGGAAATTCGCATGGTCAACATGAAAGTAAGGAACCTCCCGGGGATCTAGATTCTTC	3226
Qy	54	SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu	73
Db	3225	TCTAGATTAGTAGTTGGAGAAATTCACAGTGNCTGGACATCCATGGCAGGCTCCCTA	3166
Qy	74	LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr	93
Db	3165	AAATCAGATGAGCACCACCTCTCTGTGGAGGAAGCTTGATTCAAGAAGATCGGGTGTGTACA	3106
Qy	94	AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThra	111
Db	3105	GCAGCACATCGCTGGACACGCCCTCAGTGAAGCAGCTGAAGATATAACTGTGACTCT	3046
Qy	112	GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	131
Db	3045	GGGGAGTACAGCTCTTTCAGAGGATAACCAAGACAGATATCTCTGTCTCAAAATTT	2986
Qy	132	IleIleHisProHisSerThrLysSerProMetAspTyrAspIleAlaLeuLeuLys	151
Db	2985	ATTACCCATCCTGAATACACAGCCGGTAATATATAGTCTCTGATATGCTACTCTGTAT	2926
Qy	152	MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171
Db	2925	CTAAACACAAAGTCAAGTGTGGAAATGCTGTTGAGCCAAATCTGCTCTCTGACAGCGAT	2866
Qy	172	GluGlnPheGluAlaGlyPheIleCysThrThraIalaGlyTrpGlyArgLeuThrGluGly	191

Db	2865	GATAAAGTTGACACGAGAAATTCCTTGTCTATCCAGTGGTGGCAAGATTTCACAAACA	2806
Qy	192	GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys	211
Db	2805	TCAGAATATTCAAAATGCTCTACAGAATAATGGAATTCCTCCATCATGATGACAGACGCTGT	2746
Qy	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231
Db	2745	AATACTGTCTCAAGACGATGAACCTCCCTCCCTGGAGGACCATGCTGTGTCTGGC	2686
Qy	232	PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251
Db	2685	TTCCCTGATGGGAATGGAGCGCTGCCAGGGGACTCTGGAGGACCATGGTTGTGTAGA	2626
Qy	252	AsnLysLysGlyAlaIleThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg	271
Db	2625	AGAGGTGGTGAATCTCGATTCTTGTGGATAACTTCTCTGGGTAGCTGGTGTGCTGGA	2566
Qy	272	Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe	288
Db	2565	GGTTCAGTTCCGTAAAGAACCAACCATGTGAAGCA-----TCACTTGGCATTTTC	2515
Qy	289	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly	305
Db	2514	TCCAAAGTGTCTGATGGATTGATTTATCACTCAAAACCTGTTTCACAGGT	2464

RESULT 10

AAS72057
 ID AAS72057 standard; cDNA; 4729 BP.

XX
 AC AAS72057;

XX
 DT 13-FEB-2002 (first entry)

XX
 DE DNA encoding novel human diagnostic protein #7861.

XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX
 OS Homo sapiens.

XX
 PN WO200175067-A2.

XX
 PD 11-OCT-2001.

XX
 PF 30-MAR-2001; 2001WO-US08631.

XX
 PR 31-MAR-2000; 2000US-0540217.

XX
 PR 23-AUG-2000; 2000US-0649167.

XX
 PA (HYSE-) HYSEQ INC.

XX
 XX Drmanac RT, Liu C, Tang YT;

XX
 PI WPI: 2001-639362/73.

XX
 DR P-PSDB; ABG07870.

XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1; SEQ ID No 7861; 103pp; English.

XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Alignment Scores:
 Pred. No.: 8.89e-45 Length: 4729
 Score: 567.50 Matches: 114
 Percent Similarity: 59.57% Conservative: 51
 Best Local Similarity: 41.16% Mismatches: 101
 Query Match: 34.69% Indels: 11
 DB: 23 Gaps: 4

US-09-735-713A-2 (1-306) x AAS72057 (1-4729)

Qy 37 CysGlyGlnSerLeuValVal-----GlnProTrpAsnTyrPheAsnIlePhe 53
 Db 2542 TGTGGAATTCGCATGCTCAACATGAAAGTAAGAAACCTGCCGTGGGATCAGATTCCTC 2601
 Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
 Db 2602 TCTAGAATTAGTAGTTGGAGAAATTCACACAGTGCAGTGCATCCATGGCAGTCTCCCTA 2661
 Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
 Db 2662 AAATCAGATGACACCATCTCTGTGGAGGAGCTTGATTCAGGAAGATCGGGTGTTCACA 2721
 Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
 Db 2722 GCAGCACACTGCTGCAGACGCTCAGTGCAGAGCAGCTGAGAAATATAACTGTGACTTCT 2781
 Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
 Db 2782 GGGGAGTACGCCTCTTCAGAGGATAAGCAAGAACAGATATATTCCTGCTCAAAAAT 2841
 Qy 132 IleIleHisProHisSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
 Db 2842 ATTACCCATCTGATACACAGCCGCTGAATATATGACTCCTGATATTGCATGCTGTAT 2901
 Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
 Db 2902 CTAAACACAAAGTCAAGTTGGAAATGCTGTTACGCCAATCTGCTTCCTGACAGCGAT 2961
 Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
 Db 2962 GATAAAGTTGAACCAAGAAATCTTTGCTTATCCAGTGTAGGGGCAAGATTTCACAAACA 3021
 Qy 192 GlyValLeuSerGlnValGlnGlnValAsnLeuProIleLeuThrTrpGluGluCys 211
 Db 3022 TCAGATATATCAATGCTCTACAGAATGGAATTCCTCCATCATGGATGACAGACGCTGT 3081
 Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
 Db 3082 AATACTGTGCTCAACAGCATGAACCTCCCTCCCTGGGAAGGACCATGCTGTGCTGGC 3141
 Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
 Db 3142 TTCCCTGTATGGGAATGGAGCCCTGCCAGGGGGGACTCTGGAGGACCACCTGTTGTAGA 3201
 Qy 252 AsnLysLysGlyAlaGlyThrThrAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
 Db 3202 AGAGTGTGGNAATCGAATCTTCTGCGGATAAATCTTCTGGTAGCTGTGTGCTCGA 3261

Qy 272 Gly-----TirArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
 Db 3262 GGTTCAGTCCCGTAAAGAACCAACCATGTGAAGCA-----TCATTCGCATTTTC 3312
 Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
 Db 3313 TCCAAAGTGTCTGAGTGTGATGATTTATCTACTCAAAACCTGTTCCACAGGT 3363

RESULT 11

AAS74405

ID AAS74405 standard; cDNA; 4729 BP.

XX AC AAS74405;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #10209.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG10218.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity -

XX PS Claim 1; SEQ ID No 10209; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful in medical
 imaging of sites expressing (II). (I) and (II) are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. AAS64197-AAS94564 represent novel human
 diagnostic coding sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Alignment Scores:

Pred. No.: 8.89e-45 Length: 4729

Score:	567.50	Matches:	114
Percent Similarity:	59.57%	Conservative:	51
Best Local Similarity:	41.16%	Mismatches:	101
Query Match:	34.69%	Indels:	11
DB:	23	Gaps:	4

US-09-735-713A-2 (1-306) x AAS74405 (1-4729)

QY	37	CysGlyClnSerLeuVal	-----GlnProTrpAsnTyrPheAsnIlePhe	53
DB	2542	TGTGGAATTCGCATGTCACATGAAGAACTGCGGTGGATCTAGATCTTC	2601	
QY	54	SerArgIleLeuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu	73	
DB	2602	TCTAGAATTAGTAGTGGAGAAATCAACAGTGAAGTGGACATCCATGGCAGGTCTCCCTA	2661	
QY	74	LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTrpValIleThr	93	
DB	2662	AAATCAGATGAGCACCCTCTGTGGAGAGCTTGATTCAGAAAGATCGGCTGTGTACA	2721	
QY	94	AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla	111	
DB	2722	GCAGCAGCTGCTGCAGACGCTCAGTGAGACGCTGAGAAATATACGTGTGACTCT	2781	
QY	112	GlyCyluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	131	
DB	2782	GGGGAGTACAGCCTCTTTCAGAGGATAAGCAAGACAGAAATATTCCTGCTCAAAAATT	2841	
QY	132	IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys	151	
DB	2842	ATTACCATCTCGAATACAAACAGCCGTGAATATATAGTCTCTGATATTGCTGTGTAT	2901	
QY	152	MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171	
DB	2902	CTAAACACAACTCAAGTTTGGAAATGCTGTACGCCAATCTCTTCCCTGACAGCCAT	2961	
QY	172	GluClnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	191	
DB	2962	GATAAAGTTGAACCCAGGAATCTCTTCTTATCCAGTGGATGGGCAAGATTTCACAAACA	3021	
QY	192	GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluCys	211	
DB	3022	TCAGAATATTCAAATCTCTACAAAGAAATGGAACTTCCCATCATGGATGGACAGACGCTG	3081	
QY	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231	
DB	3082	AATACTGTGCTCAAGAGCATGAACCTCCCTCCCTGGGAAGGACCATGCTGTGCTGGC	3141	
QY	232	PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251	
DB	3142	TTCCCTGATTGGGAATGACGCTCCAGGGGGACTCTGGAGGACCATCGTGTGTGTAGA	3201	
QY	252	AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg	271	
DB	3202	AGAGTGGTGGAAATCTGATCTTCTGCTGGATTAATCTTCTGCTGCTGCTGCTGCTGGA	3261	
QY	272	Gly-----TrpArgAsnValArgLysSerAspGlnGlySerProGlyIlePhe	288	
DB	3262	GTTTCAGTTCCTGATAGAACACACCATGATGTAAGGCA-----TCACTTGGCATTTTC	3312	
QY	289	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleClnThrGly	305	
DB	3313	TCCAAAGTGTCTGAGTTGATGGATTTATCACTCAAAACCTCTTTCACAGGT	3363	

RESULT 12
AAS78775
ID AAS78775 standard; cDNA; 4729 BP.
XX
AC
AC
XX
DT
XX
DE

DNA encoding novel human diagnostic protein #14579.

13-FEB-2002 (first entry)

AAS78775;

QY 37 CysGlyClnSerLeuVal-----GlnProTrpAsnTyrPheAsnIlePhe 53
DB 2542 TGTGGAATTCGCATGTCACATGAAGAACTGCGGTGGATCTAGATCTTC 2601
QY 54 SerArgIleLeuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
DB 2602 TCTAGAATTAGTAGTGGAGAAATCAACAGTGAAGTGGACATCCATGGCAGGTCTCCCTA 2661
QY 74 LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTrpValIleThr 93
DB 2662 AAATCAGATGAGCACCCTCTGTGGAGAGCTTGATTCAGAAAGATCGGCTGTGTACA 2721
QY 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
DB 2722 GCAGCAGCTGCTGCAGACGCTCAGTGAGACGCTGAGAAATATACGTGTGACTCT 2781
QY 112 GlyCyluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
DB 2782 GGGGAGTACAGCCTCTTTCAGAGGATAAGCAAGACAGAAATATTCCTGCTCAAAAATT 2841
QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys 151
DB 2842 ATTACCATCTCGAATACAAACAGCCGTGAATATATAGTCTCTGATATTGCTGTGTAT 2901
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
DB 2902 CTAAACACAACTCAAGTTTGGAAATGCTGTACGCCAATCTCTTCCCTGACAGCCAT 2961
QY 172 GluClnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
DB 2962 GATAAAGTTGAACCCAGGAATCTCTTCTTATCCAGTGGATGGGCAAGATTTCACAAACA 3021
QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluCys 211
DB 3022 TCAGAATATTCAAATCTCTACAAAGAAATGGAACTTCCCATCATGGATGGACAGACGCTG 3081
QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
DB 3082 AATACTGTGCTCAAGAGCATGAACCTCCCTCCCTGGGAAGGACCATGCTGTGCTGGC 3141
QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
DB 3142 TTCCCTGATTGGGAATGACGCTCCAGGGGGACTCTGGAGGACCATCGTGTGTGTAGA 3201
QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
DB 3202 AGAGTGGTGGAAATCTGATCTTCTGCTGGATTAATCTTCTGCTGCTGCTGCTGCTGGA 3261
QY 272 Gly-----TrpArgAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
DB 3262 GTTTCAGTTCCTGATAGAACACACCATGATGTAAGGCA-----TCACTTGGCATTTTC 3312
QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleClnThrGly 305
DB 3313 TCCAAAGTGTCTGAGTTGATGGATTTATCACTCAAAACCTCTTTCACAGGT 3363

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss. Homo sapiens.

WO200175067-A2.
11-OCT-2001.
30-MAR-2001; 2001WO-US08631.
31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
(HYSE-) HYSEQ INC.
Drmanac RT, Liu C, Tang YT;
WPI; 2001-639362/73.
P-PSDB; ABG14588.
New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
Claim 1; SEQ ID NO 14579; 103pp; English.
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Alignment Scores:
Pred. No.: 8.89e-45 Length: 4729
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.69% Indels: 11
DB: 23 Gaps: 4

US-09-735-7

```
Db 2662 AAATCAGATGACACCACTTCTGTGGAGGAAGCTTGATTCAAGAAAGATCGGTTGTTTACA 2721
Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
|||||
Db 2722 GCAGCACACTGCTGGACAGCTCAGTGTGAGAGCAGCTGAAGAATATTAAGTGTGACTTCT 2781
Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
|||||
Db 2782 GGGAGTACAGCTCTTTTCAGAGGATAAGCAAGAACAGAAATATTCCTGTCTCAAAAAT 2841
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
|||||
Db 2842 ATTACCCATCTCGAATACACAGCGCTGAATATATAGTCTGATTCGATTCGATGAT 2901
Qy 152 MetalGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
|||||
Db 2902 CTAAACACAAAGTCAAGTTGGAAATCTGTTACGCCAAATCTGCTTCCTGACAGCGAT 2961
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThrGluGly 191
|||||
Db 2962 GATAAAGTTGAACCAAGGAATTTTCTGCTATCCAGTGGATGGGCAAGATTTCCAAACA 3021
Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
|||||
Db 3022 TCAGATATTCGAATGCTCTACAGAAATGGAATTCCTCCATCATGGATGACAGACGGTGT 3081
Qy 212 ValAlaAlaLeuLeuThrLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
|||||
Db 3082 AATAGTGTCTCAAGCAGATGAACCTCCCTCCCTGGGAGGACCATGCTGTGCTCTGGC 3141
Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
|||||
Db 3142 TTCCCTGATTGGGGAATGGACGCCCTGCCAGGGGGACTCTGGAGGACCATGCTGTTGTAGA 3201
Qy 252 AsnLysLysGlyAlaThrThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
|||||
Db 3202 ACAGTGTGGTAATCTGGATCTTCTGGGATAAATCTCTGGTGTAGTGTGCTCTGGA 3261
Qy 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
|||||
Db 3262 GTTTCAGTCCCGTGAAGAACCAACCATGTGAAGGCA-----TCACTTGGCATTTC 3312
Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
|||||
Db 3313 TCCAAAGTGTCTGAGTTGATGATTTTATCATCTCAAAACCTGTTTCACAGGT 3363
RESULT 13
ID AAS84074
AC AAS84074;
XX AAS84074;
DE DNA encoding novel human diagnostic protein #19878.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR
XX 23-AUG-2000; 2000US-0649167.
XX (HYSEQ-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
```

```
XX WPI; 2001-639362/73.
DR P-PSDB; ABG19887.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
Pt diagnostics, forensics, gene mapping, identification of mutations
Pt responsible for genetic disorders or other traits and to assess
Pt biodiversity.
XX
XX Claim 1; SEQ ID No 19878; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 8.89e-45 Length: 4729
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.69% Indels: 11
DB: 23 Gaps: 4
```

US-09-735-713A-2 (1-306) x AAS84074 (1-4729)

```
Qy 37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTyrPheAsnIlePhe 53
|||||
Db 2542 TGTGGAATTCGATCGTCATGTAAGAAAGTAAGGAACCTGCGTGGGATCTAGATCTTC 2601
|||||
Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
|||||
Db 2602 TCTAGAAATTAGTAGTTGGAGAAATTCACAGTGACTGGACATCCATGCGAGTCTCCCTA 2661
|||||
Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
|||||
Db 2662 AAATCAGATGACACCACTTCTGTGGAGGAGCTTGATTCAAGAAAGATCGGTTGTTTACA 2721
|||||
Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
|||||
Db 2722 GCAGCACACTGCTGGACAGCTCAGTGTGAGAGGAGCTGAAGAATATTAAGTGTGACTTCT 2781
|||||
Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
|||||
Db 2782 GGGAGTACAGCTCTTTTCAGAGGATAAGCAAGAACAGAAATATTCCTGTCTCAAAAAT 2841
|||||
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
|||||
Db 2842 ATTACCCATCTCGAATACACAGCGCTGAATATATAGTCTGATTCGATTCGATGAT 2901
|||||
Qy 152 MetalGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
|||||
Db 2902 CTAAACACAAAGTCAAGTTGGAAATCTGTTTCACAGCAATCTGCTTCCTGACAGCGAT 2961
```


QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
 Db 2962 GATAAAGTTCAACCCAGCAATCTTCTGTTATCCAGTGGATGGGCAAGATTTCCAAACA 3021
 QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
 Db 3022 TCAGAAATATTCAAAATGCTCTACAGAAATGGAATGGAATCTCCCATCATGGATGACAGAGCGTGT 3081
 QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
 Db 3082 AATACTGCTCAAGACATGAACCTCCCTCCCTGGAGAGACCATGCTGTGTGCTGGC 3141
 QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
 Db 3142 TTCCTTGATTTGGGAATGGAGCGCTGCCAGGGGCACTCTGGAGGACCACTGGTTGTGTAGA 3201
 QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
 Db 3202 AGAGTGGTGAATCTGGATCTCTGTGGGATAACTTCTGGGTAGCTGGTGTGCTGGA 3261
 QY 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
 Db 3262 GGTTCAGTCCCTAGAAACACCATGTCAAGGCA-----TCACTTGGCATTTTC 3312
 QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
 Db 3313 TCCAAAGTGTCTGAGTTGATGGATTTTATCACTCAAAACCTGTTCACAGGT 3363
 RESULT 14
 ABQ86158
 ID ABQ86158 standard; DNA; 3096 BP.
 AC ABQ86158;
 XX
 XX
 DT 10-SEP-2002 (first entry)
 XX
 DE Novel human gene. SEQ ID 29.
 XX
 KW Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;
 KW nootropic; neuroprotective; immunosuppressive; haemostatic;
 KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;
 KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
 KW wound healing disorders; atherosclerosis; Parkinson's disease;
 KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 KW inflammation; neoplastic disease; nervous system disorder;
 KW cardiovascular disorders; pancreatitis; respiratory disorder;
 KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
 KW developmental abnormality; gastrointestinal ulceration; neuropathy;
 KW thyroid disorder; hypothyroidism; brain damage; colitis;
 KW cone photo- transduction deficiency; neurological disease; stroke;
 KW anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
 KW growth abnormality; precocious puberty; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200250105-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 17-DEC-2001; 2001WO-US49232.
 XX
 PR 19-DEC-2000; 2000US-256710P.
 PR 20-DEC-2000; 2000US-257048P.
 PR 09-JAN-2001; 2001US-260482P.
 PR 30-JAN-2001; 2001US-264922P.
 PR 06-FEB-2001; 2001US-266797P.
 PR 19-MAR-2001; 2001US-276988P.
 PR 04-APR-2001; 2001US-281535P.
 PR 08-MAY-2001; 2001US-289622P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.
 (GLAX) GLAXO GROUP LTD.
 XX
 XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;
 PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
 PT WPI; 2002-508784/54.
 DR P-PSDB; ABP60993.
 XX
 XX Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune
 PT disorder -
 XX
 PS Claim 2(a); Page 245-246; 335pp; English.
 XX
 CC The invention relates to an isolated polypeptide with signal sequences
 CC which allow it to be secreted extracellularly or membrane associated.
 CC The activity of polypeptides of the invention may be described as,
 CC cytostatic, vulnary, antiarteriosclerotic, antiparkinsonian, nootropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease,
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder,
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, anglogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
 CC of the invention.
 XX
 SQ Sequence 3096 BP; 890 A; 619 C; 742 G; 845 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1.95e-44 Length: 3096
 Score: 561.50 Matches: 113
 Percent Similarity: 59.42% Conservative: 51
 Best Local Similarity: 40.94% Mismatches: 101
 Query Match: 34.32% Indels: 11
 DB: 24 Gaps: 4
 US-09-735-713A-2 (1-306) x ABQ86158 (1-3096)
 QY 37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTrpPheAsnIlePhe 53
 Db 25 TGTGGAAATTCGCATGTCACATCAAGAAAGTAAGGAACCTGCCCTGGGATCTAGATCTTC 84
 QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTrpProTrpGlnValSerLeu 73
 Db 85 TCTAGATTAGTAGTTGGAGAAATTCACACAGTCACATCCATGGCAGGTCCTCCCTA 144
 QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
 Db 145 AAATCAGATGAGCACCACCTTCTGTGGAGGAAGCTTGATTCAAGAAGATCGGTGTGTACA 204
 QY 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
 Db 205 GCAGCACACTGCTGGACAGCCCTCAGAGACACACTGAAGAATAATAACTGTGACTTCT 264
 QY 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
 Db 265 GGGAGTACAGCCTCTTTTCAAGAGGATACAGACAGACAGAAATATCTCTGTCTCAAAAT 334
 QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTrpAspIleAlaLeuLeuLys 151

Db	325	ATTACCATCTGTAATCAACAGCGGTGAATATATGAGTCTCTGATATTGCACTGCTGTAT	384	PR	27-MAR-1998;	98US-0079786.
Qy	152	MetalGlyAlaPheGlnPheHisPheValGlyProIleCysLeuProGluLeuArg	171	PR	30-MAR-1998;	98US-0079920.
Db	385	CTAAACACAAAGTCTGGAAATGCTGTTACAGCAATCTGCTTCCTGACAGCGAT	444	PR	31-MAR-1998;	98US-0080105.
Qy	172	GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	191	PR	31-MAR-1998;	98US-0080107.
Db	445	GATAAAGTTGAACACGAGGAATCTTGTGCTTATCCAGTGGATGGGCAAGATTTCCAAACA	504	PR	31-MAR-1998;	98US-0080165.
Qy	192	GlyValLeuSerGlnValGlnGluValAsnLeuProIleLeuThrTrpGluGluCys	211	PR	31-MAR-1998;	98US-0080194.
Db	505	TCAGAAATATTCAAATGCTCTACAGAAATGGAACCTCCCATCATGGATGACAGACGGTGT	564	PR	01-APR-1998;	98US-0080327.
Qy	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231	PR	01-APR-1998;	98US-0080333.
Db	565	AATACTGTGCTCAAGACATGAACCTCCCTCCCTGGGAAGACCATGCTGTGCTGGC	624	PR	01-APR-1998;	98US-0080334.
Qy	232	PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251	PR	08-APR-1998;	98US-0081049.
Db	625	TTCCTGTATTGGGAATGAGCCCTGCCAGGGGACTCTGGAGGACCACTGGTTGTAGA	684	PR	08-APR-1998;	98US-0081071.
Qy	252	AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg	271	PR	09-APR-1998;	98US-0081195.
Db	685	AGAGGTGTGGAATCTGGATTCTTCTGGGATACTTCTCGGTAGCTGGTTGCTGGA	744	PR	09-APR-1998;	98US-0081203.
Qy	272	Gly-----TTPArgAsnValArgLysSerAspGlnGlySerProGlyIlePhe	288	PR	09-APR-1998;	98US-0081229.
Db	745	GGTTCAGTTCCTGTAAGAAACCAACATGTGAAGCA-----TCACTGGCATTTTC	795	PR	15-APR-1998;	98US-0081817.
Qy	289	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr	304	PR	15-APR-1998;	98US-0081838.
Db	796	TCCAAAGTGTCTGATTCATGATTTTATCACTCAAAACCTGTTTCAACA	843	PR	15-APR-1998;	98US-0081952.
RESULT 15						
AAZ34034						
ID	AAZ34034 standard; DNA; 1327 BP.					
XX	AAZ34034;					
XX	AAZ34034;					
DT	07-DEC-1999 (first entry)					
XX	XX					
DE	Human EST DNA35597.					
XX	XX					
KW	Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;					
KW	probe; blood coagulation disorder; cancer; cellular adhesion disorder;					
XX	secreted protein; transmembrane protein; ss.					
OS	Homo sapiens.					
XX	XX					
PN	W09946281-A2.					
XX	XX					
PD	16-SEP-1999.					
XX	XX					
PF	08-MAR-1999; 99WO-US05028.					
XX	XX					
PR	10-MAR-1998; 98US-0077450.					
PR	11-MAR-1998; 98US-0077632.					
PR	11-MAR-1998; 98US-0077641.					
PR	11-MAR-1998; 98US-0077649.					
PR	12-MAR-1998; 98US-0077791.					
PR	13-MAR-1998; 98US-0078004.					
PR	17-MAR-1998; 98US-0040220.					
PR	20-MAR-1998; 98US-0078886.					
PR	20-MAR-1998; 98US-0078910.					
PR	20-MAR-1998; 98US-0078936.					
PR	20-MAR-1998; 98US-0078939.					
PR	25-MAR-1998; 98US-0079294.					
PR	26-MAR-1998; 98US-0079656.					
PR	27-MAR-1998; 98US-0079663.					
PR	27-MAR-1998; 98US-0079664.					
PR	27-MAR-1998; 98US-0079689.					
PR	27-MAR-1998; 98US-0079728.					
PR	27-MAR-1998; 98US-0079786.					
PR	30-MAR-1998; 98US-0079920.					
PR	31-MAR-1998; 98US-0080105.					
PR	31-MAR-1998; 98US-0080107.					
PR	31-MAR-1998; 98US-0080165.					
PR	31-MAR-1998; 98US-0080194.					
PR	01-APR-1998; 98US-0080327.					
PR	01-APR-1998; 98US-0080333.					
PR	01-APR-1998; 98US-0080334.					
PR	08-APR-1998; 98US-0081049.					
PR	08-APR-1998; 98US-0081070.					
PR	08-APR-1998; 98US-0081071.					
PR	09-APR-1998; 98US-0081195.					
PR	09-APR-1998; 98US-0081203.					
PR	09-APR-1998; 98US-0081229.					
PR	15-APR-1998; 98US-0081817.					
PR	15-APR-1998; 98US-0081838.					
PR	15-APR-1998; 98US-0081952.					
PR	15-APR-1998; 98US-0081955.					
PR	21-APR-1998; 98US-0082568.					
PR	21-APR-1998; 98US-0082569.					
PR	22-APR-1998; 98US-0082700.					
PR	22-APR-1998; 98US-0082704.					
PR	22-APR-1998; 98US-0082804.					
PR	23-APR-1998; 98US-0082767.					
PR	23-APR-1998; 98US-0082796.					
PR	27-APR-1998; 98US-0083336.					
PR	28-APR-1998; 98US-0083322.					
PR	29-APR-1998; 98US-0083392.					
PR	29-APR-1998; 98US-0083495.					
PR	29-APR-1998; 98US-0083496.					
PR	29-APR-1998; 98US-0083499.					
PR	29-APR-1998; 98US-0083500.					
PR	29-APR-1998; 98US-0083545.					
PR	29-APR-1998; 98US-0083554.					
PR	29-APR-1998; 98US-0083558.					
PR	30-APR-1998; 98US-0083559.					
PR	30-APR-1998; 98US-0083742.					
PR	06-MAY-1998; 98US-0084366.					
PR	06-MAY-1998; 98US-0084414.					
PR	06-MAY-1998; 98US-0084441.					
PR	07-MAY-1998; 98US-0084598.					
PR	07-MAY-1998; 98US-0084600.					
PR	07-MAY-1998; 98US-0084637.					
PR	07-MAY-1998; 98US-0084639.					
PR	07-MAY-1998; 98US-0084640.					
PR	07-MAY-1998; 98US-0084643.					
PR	13-MAY-1998; 98US-0085323.					
PR	13-MAY-1998; 98US-0085338.					
PR	15-MAY-1998; 98US-0085373.					
PR	15-MAY-1998; 98US-0085573.					
PR	15-MAY-1998; 98US-0085579.					
PR	15-MAY-1998; 98US-0085580.					
PR	15-MAY-1998; 98US-0085582.					
PR	15-MAY-1998; 98US-0085689.					
PR	15-MAY-1998; 98US-0085697.					
PR	15-MAY-1998; 98US-0085700.					
PR	15-MAY-1998; 98US-0085704.					
PR	18-MAY-1998; 98US-0086023.					
PR	22-MAY-1998; 98US-0086392.					
PR	22-MAY-1998; 98US-0086414.					
PR	22-MAY-1998; 98US-0086430.					
PR	22-MAY-1998; 98US-0086486.					
PR	28-MAY-1998; 98US-0087098.					
PR	28-MAY-1998; 98US-0087106.					
PR	28-MAY-1998; 98US-0087208.					
PR	30-JUL-1998; 98US-0094651.					
PR	11-SEP-1998; 98US-0100038.					
PA	(GETH) GENENTECH INC.					

(GETH) GENENTECH INC.

XX

PA

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 25, 2003, 06:30:40 ; Search time 20.5 seconds

(without alignments)

3726.800 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 1668

Sequence: 1 atgagctcacaagtgttat.....acatccaaactgtaactaa 921

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame_n2p.model -DEV-xlh
-Q/cgna2_1/USPTO_spoil/US09735713/runat_18032003_124354_23119/app_query.fasta_1.1095
-DB-SwissProt_40 -QFMT-fastan -SUFFIX-rsp -MINMATCH=0.1 -LOOPFCI=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09735713 @CGN_1_11 @runat_18032003_124354_23119 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486.5	29.2	786	1 STUB_DROME	Q05319 drosophila
2	466	27.9	638	1 KAL_HUMAN	P03952 homo sapien
3	464	27.8	638	1 KAL_MOUSE	P26262 mus musculu
4	456.5	27.4	454	1 TMS3_HUMAN	P57727 homo sapien
5	448	26.9	342	1 PSS8_MOUSE	Q9es87 rattus norv
6	446	26.7	342	1 PSS8_MOUSE	Q9esd1 mus musculu
7	446	26.7	638	1 KAL_RAT	P14272 rattus norv
8	442.5	26.5	1035	1 ENTK_BOVIN	P98072 bos taurus
9	441	26.4	343	1 PSS8_HUMAN	Q16651 homo sapien
10	440.5	26.4	1034	1 ENTK_PIG	P98074 sus scrofa
11	440	26.4	416	1 HEP5_MOUSE	Q35453 mus musculu
12	440	26.4	417	1 HEP5_HUMAN	P05981 homo sapien
13	438	26.3	416	1 HEP5_RAT	Q05511 rattus norv
14	438	26.3	457	1 TMS5_HUMAN	Q9h3s3 homo sapien
15	433	26.0	593	1 FA12_BOVIN	P98140 bos taurus
16	431.5	25.9	1019	1 FA12_HUMAN	P98073 homo sapien
17	427	25.6	311	1 TRYG_MOUSE	Q9ql77 mus musculu
18	426	25.5	855	1 ST14_MOUSE	P56677 mus musculu

19	425.5	25.5	290	1 MPN_HUMAN	Q9bqr3 homo sapien
20	425	25.5	1069	1 ENTK_MOUSE	P97435 mus musculu
21	423	25.4	455	1 TMS5_MOUSE	Q9er04 mus musculu
22	421	25.2	317	1 BSS4_HUMAN	Q9gzn4 homo sapien
23	421	25.2	321	1 TRYG_HUMAN	Q9nrr2 homo sapien
24	419.5	25.1	237	1 TRYG_ASTFL	P00765 astacus flu
25	419	25.1	264	1 CTRL_HUMAN	P40313 homo sapien
26	419	25.1	273	1 TRYT_SHEEP	Q9xsm2 ovis aries
27	415.5	24.9	275	1 TRY3_ANOGA	P35037 anopheles g
28	415	24.9	306	1 BSS4_MOUSE	Q9er10 mus musculu
29	415	24.9	625	1 FA11_HUMAN	P03951 homo sapien
30	414.5	24.9	615	1 FA12_HUMAN	P00748 homo sapien
31	413.5	24.8	263	1 CTR2_CANFA	P04813 canis famil
32	413.5	24.8	274	1 TRY1_ANOGA	P35035 anopheles g
33	413	24.8	437	1 TMS4_HUMAN	Q9nrs4 homo sapien
34	409	24.5	263	1 CTR8_HUMAN	P17538 homo sapien
35	409	24.5	490	1 TMS2_MOUSE	Q9j1q8 mus musculu
36	408.5	24.5	415	1 ACRO_PIG	P08001 sus scrofa
37	408	24.5	275	1 TRY4_ANOGA	P35038 anopheles g
38	407.5	24.4	459	1 PRTC_PIG	Q9g1p2 sus scrofa
39	407	24.4	275	1 TRYT_PIG	Q9n2d1 sus scrofa
40	406	24.3	254	1 TRY3_AEDAE	P29786 aedes aegyp
41	406	24.3	492	1 TMS2_HUMAN	O15393 homo sapien
42	405.5	24.3	603	1 FA12_CAVPO	Q04962 cavia porce
43	404	24.2	407	1 FA7_BOVIN	P22457 bos taurus
44	403	24.2	314	1 TEST_HUMAN	Q9y6m0 homo sapien
45	403	24.2	324	1 TEST_MOUSE	Q9jh17 mus musculu

ALIGNMENTS

RESULT 1

ID	STUB_DROME	STANDARD;	PRT;	786 AA.
AC	Q05319;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Serine proteinase stubble (EC 3.4.21.-) (stubble-stubblid protein).			
GN	SB OR SB-SB.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Oregon-R;			
RX	MEDLINE=93281671; PubMed=7685111;			
RA	Appel L.F., Prout M., Abu-Shumays R., Hammonds A., Garbe J.C.,			
RA	Fristrom D., Fristrom J.			
RT	"The Drosophila Stubble-stubblid gene encodes an apparent			
RT	transmembrane serine protease required for epithelial			
RT	morphogenesis."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:4937-4941(1993).			
CC	-I- FUNCTION: HORMONE DEPENDENT PROTEASE REQUIRED FOR EPITHELIAL			
CC	MORPHOGENESIS. HAS A DUAL FUNCTION, DETACHES IMAGINAL DISC CELLS			
CC	FORM EXTRACELLULAR MATRICES THROUGH ITS EXTRACELLULAR PROTEOLYTIC			
CC	DOMAIN AND TRANSMITS AN OUTSIDE-TO-INSIDE SIGNAL TO ITS			
CC	INTRACELLULAR DOMAIN TO MODIFY THE CYTOSKELETON DURING			
CC	MORPHOGENESIS. MAY BE ABLE TO ACTIVATE ITSELF.			
CC	-I- SUBCELLULAR LOCATION: Type II membrane protein.			
CC	-I- INDUCTION: BY 20-HYDROXYCYDSONE (20HE).			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-I- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-24 IS THE INITIATOR.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

```

CC  EMBL; L11451; AAA28918.1; -
DR  HSP; P00763; IDPO.
DR  MEROPS; S01.225;
DR  FlyBase; Fgn0003319; Sb.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR001254; Ser_protease_Try.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  SMART; SM00020; Tryp_SPC; 1.
DR  PROSITE; PS00240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
DR  Hydrolase; Serine protease; Transmembrane; Glycoprotein; Zymogen;
KW  Signal-anchor.
FT  CHAIN 1 542 NON-CATALYTIC CHAIN (POTENTIAL).
FT  CHAIN 543 786 CATALYTIC CHAIN (POTENTIAL).
FT  DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 59 80 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT  DOMAIN 81 786 EXTRACELLULAR (POTENTIAL).
FT  DOMAIN 543 786 SERINE PROTEASE.
FT  DOMAIN 267 276 POLY-SER.
FT  DOMAIN 287 298 POLY-GLN.
FT  DOMAIN 391 478 SER/THR-RICH.
FT  DOMAIN 412 422 POLY-THR.
FT  DOMAIN 471 478 POLY-SER.
FT  ACT_SITE 589 589 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE 639 639 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE 737 737 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  DISULFID 531 659 INTERCHAIN (BY SIMILARITY).
FT  DISULFID 574 590 BY SIMILARITY.
FT  DISULFID 703 722 BY SIMILARITY.
FT  DISULFID 733 762 BY SIMILARITY.
FT  CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ  SEQUENCE 786 AA; 85010 MW; CE3E755760B9DE4D CRC64;

```

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
25	1.54e-37	786	107
493	486.50	107	60
85	52.19%	Conservative:	102
510	33.44%	Mismatches:	51
118	29.17%	Indels:	11
530	1	Gaps:	

```

US-09-735-713a-1 (1-921) x STUB_DROME (1-786)
QY 25 AGGACAAAGCTGATTTTACTAGTAAGTAGTCTTTTGAACRAGGTAAATCTGCARCT 84
Db 493 ArgThrProValLeuAlaThrSerGlylle-----GluthrAsnGluIleSerAsp 509
QY 85 CTTTTCCTCCCAAGAGCTCCAGTGTGGGCAG-----
Db 510 SerSerIleProAspAlaGlyAlaLeuGlyArgValLysThrIleSerAlaAlaArgSer 529
QY 118 -----AGTCGTGTAGGTACAGCCTTGGAAATATTTTAAACATTTTCACT 162
Db 530 GluGlyValProThrLeuAlaArgProGlu-----Thr 541
QY 163 CGCATTTCTGGAGGAGCAAGTGGAGAGGCTTCTATCCCTGGCAGGTATCTCTGAA 222
Db 542 ArgIleValGlyLysSerAlaAlaPheGlyArgTrpProTrpGlnValSerValArg 561
QY 223 CAA-----AGGCAGAGCAGCATTTTGTGGAGGAGCATGCTCTCAACA 264
Db 562 ArgThrSerPheGlyPheSerThrHisArgCysGlyGlyAlaLeuIleAsnGlu 581
QY 265 CAGTGGGTGATCAGCGCTCAGTCATTCGACATTCGACAAACAAATGTGCTACTTTGAAT 324
Db 582 AsnTrpIleAlaThrAlaGlyHisCysValAspLeu---LeuIleSerGlnIleArg 600
QY 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATT 384
Db 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATT 384

```

```

Db 601 IleArgValGlyGluTyrAspPheSerHisValGln-----GluGlnLeuProTyrIle 618
QY 385 GAA-----ACTGTCATCATCATCATCCACATTTCTCCACCAAGAAACCAATGCAC 432
Db 619 GluArgGlyValAlaLysLysValValHisProLysTyrSerPhe---LeuThrTyrGlu 637
QY 433 TATGATATTCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTTGGGGCCC 492
Db 638 TyrAspLeuAlaLeuValLysLeuGluGlnProLeuGluPheAlaProHisValSerPro 657
QY 493 ATATGCTTTCACAGCTGCGGAGCAATTTGAGCGTGGTATTATTTGTACAACTGCAGGC 552
Db 658 IleCysLeuProGlu---ThrAspSerLeuLeuIleGlyMetAsnAlaThrValThrGly 676
QY 553 TGGGGCGGCTTAAGTGAAGTGGCTCTCATCAAGTCTTGCAGGAAGTGAATCTGCCT 612
Db 677 TrpGlyArgLeuSerGluGlyThrLeuProSerValLeuGlnGluValSerValPro 696
QY 613 ATTTTGACCTGGGAAGAGTGTGGCAGCTCTGTTAAACACTA---AAGAGGCCCATCAGT 669
Db 697 IleValSerAsnAspAsnCysLysSerMetPheMetArgAlaGlyArgGlnGluPheIle 716
QY 670 GGAAGACCTTCTTTGCACAGGTTTCTGATGGAGGAGAGACCATGTCAGGAGAT 729
Db 717 ProAspIlePheLeuGlyCysAlaGlyTyrGluThrGlyGlyGlnAspSerCysGlnGlyAsp 736
QY 730 TCAGGAGGTCTCATCTGTCGCGGAATAAGAAAGGGCGCTGGACTCTGCTGGTGTGACT 789
Db 737 SerGlyGlyProLeuGlnAlaLysSerGlnAspGlyArgPhePheLeuAlaGlyIleIle 756
QY 790 TCTGGGGTTTGGCTGTGGTCTGAGGCTGGAGAAACAATGTGAGGAAATGATCAAGGA 849
Db 757 SerTrpGlyIleGlyCys-----AlaGluAlaAsn 766
QY 850 TCCCTGGGATCTTCACACACATTTAGTAAGTCTTCTGATCCACGACACATCCAA 909
Db 767 LeuProGlyValCysThrArgIleSerLysPheThrProTrpIleLeuGluHisValArg 786

RESULT 2
KAL_HUMAN
ID KAL_HUMAN STANDARD; PRT; 638 AA.
AC P03952;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (Kininogenin) (Fletcher factor).
GN KLKB1 OR KLK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86243359; PubMed=3521732;
RA Chung D.W., Fujikawa K., McMullen B.A., Davie E.W.;
RT "Human plasma prekallikrein, a zymogen to a serine protease that
RT contains four tandem repeats.";
RL Biochemistry 25:2410-2417(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS SER-143; GLN-202 AND PRO-208.
RX MEDLINE=20487549; PubMed=11031105;
RA Yu H., Anderson P.J., Freedman B.I., Rich S.S., Bowden D.W.;
RT "Genomic structure of the human plasma prekallikrein gene,
RT identification of allelic variants, and analysis in end-stage renal
RT disease.";
RL Genomics 69:225-234(2000).
RN [3]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=91152016; PubMed=1998666;
RA McMullen B.A., Fujikawa K., Davie E.W.;
RT "Location of the disulfide bonds in human plasma prekallikrein: the
RT presence of four novel apple domains in the amino-terminal portion of

```



```

Db 548 CysGlnLysArgTyrGlnAspTyrLys-----IleThrGlnArgMetValCysAla 564
QY 691 GTTTCTCTGATGAGGAGAGAGCGATGTCAGGAGATTCAGGAGGTTTCACTCATGTGC 750
Db 565 GlyTyrLysGluGlyLysAspAlaCysLysGlyAspSerGlyGlyProLeuValCys 584
QY 751 CGAATAAGAAAGGGCGCTGGACTCTGGCTGGTGTGCTGACTTCCCTGGGGTTTGGCTGTGGT 810
Db 585 LysHis---AsnGlyMetTTPArgLeuValGlyIleThrSerTrpGlyGluGlyCysAla 603
QY 811 CGAGGCTGGAGAACATGTCAGGAAAGTATCATCAGGATCCCTGGGATCTTCACAGAC 870
Db 604 Arg-----ArgGluGlnProGlyValTyrThrLys 613
QY 871 ATTAGTAAGTCTGCTTCTCGATCCACGACACATCAAACTGGTAAC 918
Db 614 ValAlaGluTyrMetAspTrpIleLeuGluLysThrGlnSerSerAsp 629

RESULT 3
KAL_MOUSE
ID KAL_MOUSE STANDARD; PRT: 638 AA.
AC P26262;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (kininogenin) (Fletcher factor).
GN KLB1 OR KLB3 OR PK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=91090844; PubMed=2264928;
RA Seidman N.G., Sawyer N., Hamelin J., Mion P., Beaubien G.,
RA Brachpaga L., Rochmont J., Mbikay M., Chretien M.;
RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,
RT and comparison of protein and mRNA levels among species.";
RL DNA Cell Biol. 9:737-748(1990).
CC -!- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -!- CATALYTIC ACTIVITY: cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: THE ZYMOMEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 4 APPLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M58588; AAA63393.1; -.
CC PIR: A36557; KMSPL.
CC HSP: P00750; IRTF.
CC MEROPS: S01.212; -.
CC MGD: MGI:102849; Klb3.
CC InterPro: IPR000177; Apple.
CC InterPro: IPR001314; Chymotrypsin.

```

```

DR InterPro: IPR003014; PAN.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00005; APPLIEDOMAIN.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00223; APPLE; 4.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00495; APPLE; 4.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat.
FT SIGNAL 1 19 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 20 390 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 O-LINKED (PROBABLE).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (PROBABLE).
FT ACT_SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 104 BY SIMILARITY.
FT DISULFID 47 77 BY SIMILARITY.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 111 194 BY SIMILARITY.
FT DISULFID 137 166 BY SIMILARITY.
FT DISULFID 141 147 BY SIMILARITY.
FT DISULFID 201 284 BY SIMILARITY.
FT DISULFID 227 256 BY SIMILARITY.
FT DISULFID 231 237 BY SIMILARITY.
FT DISULFID 292 375 BY SIMILARITY.
FT DISULFID 318 347 BY SIMILARITY.
FT DISULFID 322 328 BY SIMILARITY.
FT DISULFID 340 345 BY SIMILARITY.
FT DISULFID 383 503 BY SIMILARITY.
FT DISULFID 419 435 BY SIMILARITY.
FT DISULFID 517 584 BY SIMILARITY.
FT DISULFID 548 563 BY SIMILARITY.
FT DISULFID 574 602 BY SIMILARITY.
SQ SEQUENCE 638 AA; 71368 MW; CC27C93AB1086599 CRC64;

Alignment Scores:
Pred. No.: 1.89e-35 Length: 638
Score: 464.00 Matches: 95
Percent Similarity: 53.49% Conservative: 43
Best Local Similarity: 36.82% Mismatches: 102
Query Match: 27.82% Indels: 18
DB: 1 Gaps: 6

US-09-735-713a-1 (1-921) x KAL_MOUSE (1-638)
QY 154 ATTTTCAGTCGCATTTTGGAGAAAGCCAAAGTGGAGAGGTTCTCTATCCCTGGCAGGTA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGlyGluTrpProTrpGlnVal 406
QY 214 TCTCTGAACAACAAAGG-----CAGAAGCATATTTTGGAGGAAGCATCTCTCACA 264
Db 407 SerLeuGlnValLysLeuValSerGlnThrHisLeuCysGlyGlySerIleIleGlyArg 426
QY 265 CAGTGGGTGATCACGGCGGCTCACGTGCATTTGCAACAGAAACATTGTGTCTACTTTGAAT 324
Db 427 GlnTrpValLeuThrAlaAlaHisCysPheAspGlyIleProTyrProAspValTrpArg 446

```



```
QY 325 GTTACTGCTGGAGATGATCACTTAAGCCAGACAGACCCAGGAGACAACTCTCACTATT 384
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 447 lletyGlyGlyIleLeuSerLeuSerGluIleThrLysGluThrProSerSerArgile 466
QY 385 GAACACTGTCATCATACATCCATCTCCACCAAGAACCAATGGACTATGATATGCC 444
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 467 LysGluLeuIleIleHisGlnLysValSerGluGly--AsnTyrAspIleAla 485
QY 445 CTTTGAAGATGGCTGGAGCCCTTCAATTTGGCCCACTTTGGGGCCCATGATCTTCCA 504
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 486 LeuIleLysLeuGlnThrProLeuAsnTyrThrGluPheGlnLysProIleCysLeuPro 505
QY 505 GAGCTGCGGGAGCAATTTGAGCGCTGTTTATTGTACAACTGCAGGCTGGGCGGCTTA 564
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 506 SerLysAlaAspThrAsnThrIleTyrThrAsnCysTyrPValThrGlyTyrThr 525
QY 565 ACTGAAGTGGCGCTCTCACAAGTCTGAGGAGTGAATCGCTATTTGACCTGG 624
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 526 LysGluGlnGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
QY 625 GAAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTCT 684
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 546 GluGluCysGlnLys-----LysTyrArgAspTyrValIleAsnLysGlnMetile 562
QY 685 TGCACAGCTTTTCTGATGAGGAGACGCATGTCAGGAGATTCAGGAGGTTCACTC 744
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 563 CysAlaGlyTyrLysGluGlyThrAspAlaCysLysGlyAspSerGlyGlyProLeu 582
QY 745 ATGTCCCGGATAGAAGGGCGCTGACTCTGCTGTGTGACATCTCTGGGTTTGGC 804
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 583 ValCysLysHis---SerGlyArgTyrGlnLeuValGlyIleThrSerTyrpGlyGly 601
QY 805 TGTGTCTCAGCTGGGAGAAACAATCTGAGGAAAGTCAAGGATCCCTCGGATCTC 864
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 602 CysGlyArg-----LysAspGln-----ProGlyValTyr 611
QY 865 ACAGACATTAGTAAGTCTCTCTGTCGATCCAGCAACATCCAACTGGTAAC 918
DB   |||  |||||:  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 612 ThrLysValSerGluTyrMetAspTyrIleLeuGluLysThrGlnSerSerAsp 629
RESULT 4
TMS3_HUMAN
ID TMS3_HUMAN STANDARD; PRT; 454 AA.
AC P57727;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 3 (EC 3.4.21.-) (Serine protease
DE TADG-12) (Tumor associated differentially-expressed gene-12 protein).
GN TMPSR3 OR TADG12 OR ECOS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
PI 1
RN RP SEQUENCE FROM N.A. (ISOFORMS A AND TRUNCATED).
RC TISSUE=Ovarian carcinoma;
RX MEDLINE=20521358; PubMed=11069177;
RA Underwood L.J., Shigemasa K., Tanimoto H., Beard J.B., Schneider E.N.,
RA Wang Y., Parmley T.H., O'Brien T.J.;
RA "Ovarian tumor cells express a novel multi-domain cell surface serine
RA protease.";
RL Biochim. Biophys. Acta 1502:337-350(2000).
RN [2]
RX SEQUENCE FROM N.A. (ISOFORMS A; B/C AND D), AND VARIANT ILE-53.
RX MEDLINE=20578749; PubMed=11137999;
RA Scott H.S., Kudoh J., Wattenhofer M., Shibuya K., Berry A., Chraast R.,
RA Guipponi M., Wang J., Kawasaki K., Asakawa S., Minoshima S.,
RA Younis F., Mehdi S.O., Radhakrishna U., Papasavvas M.P., Gehrig C.,
RA Rossier C., Korostishevsky M., Gal A., Shimizu N., Bonne-Tamir B.,
RA Antonarakis S.E.;
RA "Insertion of beta-satellite repeats identifies a transmembrane
RA protease causing both congenital and childhood onset autosomal
```

```
RT recessive deafness.";
RL Nat. Genet. 27:59-63(2001).
RN [3]
RN VARIANTS DFN8/DFNB10 CYS-251 AND LEU-404.
RX MEDLINE=21354482; PubMed=11462234;
RA Masmodi S., Antonarakis S.E., Schwede T., Ghorbel A.M., Gratri M.,
RA Papasavvas M.P., Drira M., Elgaled-Boullila A., Wattenhofer M.,
RA Rossier C., Scott H.S., Ayadi H., Guipponi M.;
RA "Novel missense mutations of TMPSR3 in two consanguineous Tunisian
RA families with non-syndromic autosomal recessive deafness.";
RL Hum. Mutat. 18:101-108(2001).
RN [4]
RN VARIANTS DFN8/DFNB10 TRP-109; PHE-194 AND ARG-407, AND VARIANTS
RN ILE-53; SER-111 AND VAL-253.
RX PubMed=11424922;
RA Ben-Yosef T., Wattenhofer M., Riazuddin S., Ahmed Z.M., Scott H.S.,
RA Kudoh J., Shibuya K., Antonarakis S.E., Bonne-rimir B.,
RA Radhakrishna U., Naz S., Ahmed Z., Riazuddin S., Pandya A.,
RA Nance W.E., Wilcox E.R., Friedman T.B., Morell R.J.;
RA "Novel mutations of TMPSR3 in four DFN8/B10 families segregating
RA congenital autosomal recessive deafness.";
RL J. Med. Genet. 38:396-400(2001).
RN [5]
RN VARIANT DFN8/DFNB10 GLY-103, AND VARIANTS ASN-173 AND THR-426.
RX PubMed=11907649;
RA Wattenhofer M., Di Iorio V., Rabionet R., Dougherty L., Pampanos A.,
RA Schwede T., Montserrat-Sentis B., Arbones L., Illades T.,
RA Pasquaddibisceglie A., D'Amelio M., Alwan S., Rossier C., Dahl H.-H.M.,
RA Petersen M.B., Estivill X., Gasparini P., Scott H.S.,
RA Antonarakis S.E.;
RA "Mutations in the TMPSR3 gene are a rare cause of childhood
RA nonsyndromic deafness in Caucasian patients.";
RL J. Mol. Med. 80:124-131(2002).
RN CC - FUNCTION: Probable protease.
RN CC - SUBCELLULAR LOCATION: Type II membrane protein (Potential).
RN CC - ALTERNATIVE PRODUCTS: 4 ISOFORMS: A (SHOWN HERE), B/C, D AND
RN CC TRUNCATED/TADG-12V; ARE PRODUCED BY ALTERNATIVE SPLICING.
RN CC - TISSUE SPECIFICITY: Expressed in many tissues including fetal
RN CC cochlea. The truncated isoform is found at increased levels in
RN CC some carcinomas.
RN CC - DISEASE: Defects in TMPSR3 are a cause of two forms of autosomal
RN CC neurosensory childhood-onset forms of deafness, DFN8 and DFNB10.
RN CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
RN CC - SIMILARITY: CONTAINS 1 SRCR DOMAIN.
RN CC - SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
RN CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RN CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
RN CC the European Bioinformatics Institute. There are no restrictions on its
RN CC use by non-profit institutions as long as its content is in no way
RN CC modified and this statement is not removed. Usage by and for commercial
RN CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN CC or send an email to license@isb-sib.ch).
```


OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX [1] NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RP MEDLINE-91129236; PubMed=1993180;
RX Beaubien G., Rosinski-Chupin I., Mattei M.-G., Mbikay M., Chretien M.,
RA Seidah N.G.;
RT "Gene structure and chromosomal localization of plasma kallikrein.";
RL Biochemistry 30:1628-1635(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-90091743; PubMed=2598771;
RA Seidah N.G., Ladenheim R., Mbikay M., Hamelin J., Lutfalla G.,
RA Rougeon F., Lazure C., Chretien M.;
RT "The cDNA structure of rat plasma kallikrein.";
RL DNA 8:563-574(1989).
CC -|- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-I-Xaa and Lys-I-Xaa
CC bonds, including Lys-I-Arg and Arg-I-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -|- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -|- SIMILARITY: CONTAINS 4 APPLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62357; AAA74563.1; -;
CC EMBL; M62358; AAA74563.1; JOINED.
CC EMBL; M62346; AAA74563.1; JOINED.
CC EMBL; M62347; AAA74563.1; JOINED.
CC EMBL; M62349; AAA74563.1; JOINED.
CC EMBL; M62350; AAA74563.1; JOINED.
CC EMBL; M62351; AAA74563.1; JOINED.
CC EMBL; M62352; AAA74563.1; JOINED.
CC EMBL; M62353; AAA74563.1; JOINED.
CC EMBL; M62354; AAA74563.1; JOINED.
CC EMBL; M62355; AAA74563.1; JOINED.
CC EMBL; M62356; AAA74563.1; JOINED.
CC EMBL; M30282; AAA41463.1; -;
CC EMBL; M58590; AAA42069.1; -;
CC PIR: A39180; KQRTPL.
CC HSSP: P00750; IRTF.
CC MEROPS: S01.212; -;
CC InterPro: IPR000177; Apple.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR003014; PAN.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam: PF00024; PAN; 4.
CC Pfam: PF00089; trypsin; 1..
CC PRINTS: PR00005; APPLIEDOMAIN.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00223; APPLE; 4.
CC SMART: SM00020; Tryp_SPC; 1.
CC PROSITE: PS00495; APPLE; 4.
CC PROSITE: PS02040; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase: Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;

KW Repeat. 1 19
FT SIGNAL 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (PROBABLE).
FT ACT_SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 104 BY SIMILARITY.
FT DISULFID 47 77 BY SIMILARITY.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 111 194 BY SIMILARITY.
FT DISULFID 137 166 BY SIMILARITY.
FT DISULFID 141 147 BY SIMILARITY.
FT DISULFID 201 284 BY SIMILARITY.
FT DISULFID 227 256 BY SIMILARITY.
FT DISULFID 231 237 BY SIMILARITY.
FT DISULFID 292 375 BY SIMILARITY.
FT DISULFID 318 347 BY SIMILARITY.
FT DISULFID 322 328 BY SIMILARITY.
FT DISULFID 340 345 BY SIMILARITY.
FT DISULFID 383 503 BY SIMILARITY.
FT DISULFID 419 435 BY SIMILARITY.
FT DISULFID 517 584 BY SIMILARITY.
FT DISULFID 548 563 BY SIMILARITY.
FT DISULFID 574 602 BY SIMILARITY.
SQ SEQUENCE 638 AA; 71273 MW; 454BEB27E8CA8F88 CRC64;

Alignment Scores:
Pred. No.: 9.14e-34 Length: 638
Score: 446.00 Matches: 91
Percent Similarity: 53.12% Conservativeness: 45
Best Local Similarity: 35.55% Mismatches: 102
Query Match: 26.74% Indels: 18
DB: 1 Gaps: 5

US-09-735-713A-1 (1-921) x KAL_RAT (1-638)

Qy 154 ATTTTCAGTCGATTCTTGGAGGAGCCAGTGGAGAGGGTTCCTATCTCCCTGGCAGGTA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnSerSerLeuGlyGluTrpProTrpGlnVal 406
Qy 214 TCTCTGAACAAGG-----CAGAAGCATATTTCTGGAGGAGCATGCTCTCACCA 264
Db 407 SerLeuGlnValLysLeuValSerGlnAsnHisMetCysGlyGlySerIleIleGlyArg 426
Qy 265 CAGTGGGTGATCACGGCGGCTCACTGTCATTCGCAACAGAAACATTTGTCTACTTTGAAT 324
Db 427 GlnTrpIleLeuThrAlaAlaHisCysPheAspGlyIleProTrpProAspValTrpArg 446
Qy 325 GTTACTGCTGGAGTAGTATGACTTAAGCCAGACAGACCAGGAGAGCAACTCTCACTATT 384
Db 447 IleTyrglyGlyIleLeuAsnLeuSerGluIleThrAsnLysThrProPheSerSerIle 466
Qy 385 GAAACTGTCATCATACATCCACATTCCTCCACCAAGAACCAATGCATGATGATATTCGC 444
Db 467 LysGluLeuIleIleHisGlnLysTyrlLysMetSerGluGly---SerTyrlAspIleAla 485
Qy 445 CTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCCA 504
Db 486 LeuIleLysLeuGlnThrProLeuAsnTyThrGluPheGlnLysProIleCysLeuPro 505
Qy 505 GAGTGGCGGGAGCAATTTGAGGCTGGTTTATTTCCTACACTGCAGCTGGGGCGGCTTA 564
Db 564
Qy 564

CC CELLS, SECRETED INTO THE DUCTS, AND EXCRETED INTO THE SEMINAL
 CC FLUID.
 CC
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC

DR EMBL; L41351; AAC41759.1; -;
 DR EMBL; U33446; AAB19071.1; -;
 DR EMBL; BC001462; AAH01462.1; -;
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.159; -;
 DR Genew; HGNC:9491; PRSS8.
 DR MIM; 600823; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP-SPC; 1.
 DR PROSITE; PS0240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;
 KW Transmembrane.
 KW SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 32 ACTIVATION PEPTIDE.
 FT CHAIN 33 44 PROSTATIN LIGHT CHAIN.
 FT CHAIN 45 322 PROSTATIN HEAVY CHAIN.
 FT PROPEP 323 343
 FT TRANSXEM 320 340
 FT DOMAIN 45 286
 FT DISULFID 37 154
 FT DISULFID 70 86
 FT DISULFID 168 244
 FT DISULFID 201 223
 FT DISULFID 234 262
 FT ACT_SITE 85 85
 FT ACT_SITE 134 134
 FT ACT_SITE 238 238
 FT CARBOHYD 159 159
 SQ SEQUENCE 343 AA; 36431 MW; 98DD6447F5A8C1B2 CRC64;
 N-LINKED (GLCNAC. . .) (POTENTIAL).

Alignment Scores:

Pred. No.:	2.35e-33	Length:	343
Score:	441.00	Matches:	100
Percent Similarity:	49.16%	Conservative:	46
Best Local Similarity:	33.67%	Mismatches:	117
Query Match:	26.44%	Indels:	34
DB:	1	Gaps:	10

US-09-735-713a-1 (1-921) x PSS8_HUMAN (1-343)

QY 34 CTGATTTTACTAGTAAGTCTTTTGTGACRAGGTAAATCTGCARCTCTTTGCTC 93
 Db 17 ILEuLeuTyrLeuGlyLeuLeuArgSerGlyThrGlyAlaGluGlyAla----- 33
 QY 94 CCCAAAGCTCCCAAGTGTGGCGACAGCTCTGGTTAAGTCAGCTTGGAAATATTTTAAAC 153
 Db 34 ---GluAlaPro---CysGly-----ValalaProGln----- 42
 QY 154 ATTTTCAGTCGCATTTTGGAGAACCCAGTGGAGAGGGTTCCTATCCCTGGCAGGTA 213
 Db 43 -----AlaargilethrGlyGlySerAlaValalaGlyGlntrpGlnVal 60
 QY 214 TCCTCTGAACAAGGAGAGCATATTTGTGAGGAGGAGCATCTCTCACCACAGTGGTG 273
 Db 61 SerilethrTyrGluGlyValHisValCysGlyGlySerLeuValSerGluGlntrpVal 80

QY 274 ATCAGCGCGCTCACTGCATTCGCAACAGAAACATTTGTCTACTTTGAATGTACTGCT 333
 Db 81 LeuSerAlaAlaHisCysPheProSerGluHisHisLysGluAlaTyrGluValLysLeu 100
 QY 334 GGAGACTATGACTTAAGCCAGACAGACCAGAGACAACTCTCACTATTGAACTGTC 393
 Db 101 GlyAlaHisGlnLeuAspSerTyrSerGluAspAlaLysValSerThrLeuLysAspIle 120
 QY 394 ATCATATCCACATTTCTCCACCAAGAACCAATGAGTATGATATTCCTCTTTGAAG 453
 Db 121 IleProHisProSerTyr---LeuGlnGluGlySerGlnGlyAspIleAlaLeuGln 139
 QY 454 ATGGCTGGAGCCCTCCCAATTTGGGCCCATATGTCTCCACAGAGTCGCG 513
 Db 140 LeuSerArgProIleThrPheSerArgTyrIleArgProIleCysLeuProAlaAlaAsn 159
 QY 514 GAGCAATTTGAGGCTGGTTTATTTGTCAACTGCAGGTGGCGCGCTTAAGTGAAGT 573
 Db 160 AlaSerPheProAsnGlyLeuHisCysThrValThrGlyTyrGlyHisValAlaProSer 179
 QY 574 GCGGTCTC-----TCACAAGTCTTGCAGGAAGTGAATCTGCCTATTTGACCTGGAA 627
 Db 180 ValSerLeuLeuThrProLysProLeuGlnGlnLeuGluValProLeuLysSerArgGlu 199
 QY 628 GAGTGTGTGGCAGCTCTGTAACTAAAGAGGCC-----ATCAGTGGGAAG 675
 Db 200 ThrCysAsnCysLeuTyrAsnIleAspAlaLysProGluGluProHisPheValGlnGlu 219
 QY 676 ACCTTTCTTTCACAGGTTTCTGTGATGAGGAGAGAGCATGTGTCAGGAGATTCAGGA 735
 Db 220 AspMetValCysAlaGlyTyrValGluGlyGlyLysAspAlaCysGlnGlyAspSerGly 239
 QY 736 GGTTCACATCTGCGCGAATAAGAAAGGCGCTGCAGCTGCTGCTGCTGCTGCTGCTG 795
 Db 240 GlyProLeuSerCys---ProValGluGlyLeuTyrTyrLeuThrGlyIleValSerTrp 258
 QY 796 GGTGTGGCTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAAGTATCAAGGATCCCC 855
 Db 259 GlyAspAlaCysGlyAla-----ArgAsnArg-----Pro 268
 QY 856 GGAATCTTCACAGACATTTAGTAAAGTCTTCTGTCGATCCAGACACATC 906
 Db 269 GlyValTyrThrLeuAlaSerSerTyrAlaSerTrpIleGlnSerLysVal 285

RESULT 10
 ENTK_PIG STANDARD; PRT; 1034 AA.
 ID ENTK_PIG
 AC P98074;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enteroproteptidase precursor (EC 3.4.21.9) (Enterokinase).
 GN PRSS7 OR ENTK.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Duodenal mucosa;
 RX MEDLINE=94327548; PubMed=8051081;
 RA Matsushima M., Ichinose M., Yahagi N., Kakei N., Tsukada S.,
 RA Miki K., Kurokawa K., Tashiro K., Shiohara K., Shinomiya K.,
 RA Uneyama H., Inoue H., Takahashi T., Takahashi K.;
 RT "Structural characterization of porcine enteroproteptidase";
 RL J. Biol. Chem. 269:19976-19982(1994).
 CC -1- FUNCTION: RESPONSIBLE FOR INITIATING
 CC PROTEOLYTIC PROENZYMES (TRYPSIN, CHYMOTRYPSIN AND CARBOXYPEPTIDASE
 CC A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN
 CC TURN ACTIVATES OTHER PROENZYMES INCLUDING CHYMOTRYPSINOGEN,
 CC PROCARBOXYPEPTIDASES, AND PROELASTASES.
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-|-Ile-7 bond in
 CC trypsinogen.


```

Db 845 ArgAsnLeuGluProSerLysTrpLysAlaIleLeuGlyLeuHisMetThrSer----- 862
QY 340 TATGACTTAAGCCAGACAGACAGGAGAGCAAACTCTCACTATTATAAATGTCATCATATA 399
Db 863 -----AsnLeuThrSerProGlnIleValThrArgLeuIleAspGluIleValle 879
QY 400 CATCACATTTCTCCACCAAGAACCAATGGACTATATATGCCCCTTTGAAGATGGCT 459
Db 880 AsnProHisTyrAsnArgArg---LysAspSerAspIleAlaMetMetHisLeuGlu 898
QY 460 GGACCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTCCAGAGCTGCGGAGCAA 519
Db 899 PheLysValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGluAsnGlnVal 918
QY 520 TTTGAGGCTGTTTATTATTGTAACACTGCAGGCTGGGCGCTTAACCTGAAGTGGCGTC 579
Db 919 PheProGlyArgIleCysSerIleAlaGlyTyrGlyValIleTyrGlnGlySer 938
QY 580 CTCTCACAGTCTTGCAGGAAGTGAATCTCCCTATTGTTGACCTGGGAGAGTGTGTGGCA 639
Db 939 ProAlaAspIleLeuGlnGluAlaAspValProLeuLeuSerAsnGluLysCysGlnGln 958
QY 640 GCTCTG-----TTAACACTAAAGAGGCCCATCATGCTGGAGACCTTCTTTCG 687
Db 959 GlnMetProGluTyrAsnIleThr-----GluAsnMetMetCys 971
QY 688 ACAGGTTTCTGATGGAGAGAGCATGTGTCAGGAGATTCAGGAGGTTTCATCTCATG 747
Db 972 AlaGlyTyrGluGluGlyGlyIleAspSerCysGlnGlyAspSerGlyGlyProLeuMet 991
QY 748 TGGCGGATTAAGAAAGGGGCTGACTCTGCTGGTGTGACTCTCTGGGGTGGGCTGT 807
Db 992 CysLeu---GluAsnAsnArgTrpLeuLeuAlaGlyValThrSerPheGlyTyrGlnCys 1010
QY 808 GGTGAGGCTGGAGAAACAATGTGAGGAAAGTGATCAAGGATCCCTGGGATCTTCACA 867
Db 1011 AlaLeuProAsnArg-----ProGlyValTyrAla 1020
QY 868 GACATTAGTAAGTGTCTTCTGATCCAGCAACATC 906
Db 1021 ArgValProLysPheThrGluTyrIleGlnSerPheLeu 1033

```

RESULT 11

```

HEPS_MOUSE STANDARD; PRT; 416 AA.
ID HEP5_MOUSE
AC 035453;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease hepsin (EC 3.4.21.-).
GN HPN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98058912; PubMed=9395459;
RA Vu T.-K.H., Liu R.W., Haakma C., Tomasek J.J., Howard E.W.;
RT Identification and cloning of the membrane-associated serine
RT protease, hepsin, from mouse preimplantation embryos.
RL J. Biol. Chem. 272:31315-31320(1997).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE
CC OF CELL MORPHOLOGY.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL; AF030065; AAB84221.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.224; -.
DR MGD; MGI:1196620; Hpn.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Sror_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolyase; Serine protease; Transmembrane; Signal-anchor.
FT CHAIN 1 161 NON-CATALYTIC CHAIN (POTENTIAL).
FT CHAIN 162 416 CATALYTIC CHAIN (POTENTIAL).
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 44 416 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 162 416 SERINE PROTEASE.
FT ACT_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 256 256 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 352 352 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 152 276 INTERCHAIN (BY SIMILARITY).
FT DISULFID 187 203 BY SIMILARITY.
FT DISULFID 321 337 BY SIMILARITY.
FT DISULFID 348 380 BY SIMILARITY.
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 416 AA; 44739 MW; 432194FF4004F848 CRC64;

```

Alignment Scores:

```

Pred. No.: 3,04e-33 Length: 416
Score: 440.00 Matches: 99
Percent Similarity: 49.47% Conservative: 41
Best Local Similarity: 34.98% Mismatches: 99
Query Match: 26.38% Indels: 44
DB: 1 Gaps: 10

```

US-09-735-713A-1 (1-921) x HEP5_MOUSE (1-416)

```

QY 109 TGTGGGCGAGAGTCTGCTTAAGGTACAGCCTTGGAATATTATTTACATTTTCAGTCGCATT 168
Db 152 CysGlyArgArgLysLeuProVal-----AspArgile 162
QY 169 CTTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228
Db 163 ValGlyGlyClnAspSerSerLeuGlyArgTrpProTrpGlnValSerLeuArgTyrAsp 182
QY 229 CAGAACCATATTTGTGGAGAACCATCTCTACACAGAGTGGGTATACAGCGGGCTCAC 288
Db 183 GlyThrHisLeuCysGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAlaAlaHis 202
QY 289 TGCATTGCAAAACAGAAAC---ATTGCTGCTACTTCTGAATGCTTACTCTGGAGAGTATGAC 345
Db 203 CysPheProGluArgAsnArgValLeuSerArgTrpArgValPheAlaAla----- 220
QY 346 TTAAGCCAGACAGACCCAGGAGAGCAAACTCTCATTATTGAAACTGTCTCATCATACATCCA 405
Db 221 ValAlaArgThrSerProHisAlaValGlnLeuGlyValGlnAlaValIleTyrHisGly 240
QY 406 CATTTTC-----TCCACCAAGAAACCAATGGAGTATGATATTGCCCTTTTG 450
Db 241 GlyTyrLeuProPheArgAspProThrIleAspGluAsnSerAsnAspIleAlaLeuVal 260
QY 451 AAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGCCCCATATGTCTTCCAGAGCTG 510
Db 261 HisLeuSerSerSerLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 280

```


QY 346 TTAAGCCAGACAGACCAGGAGAGCAAACTCTCACTATTGAACTGTCAATCATATCATCA 405
Db 222 ValAlaGlnAlaSerProHisGlyLeuGlnLeuGlyValGlnAlaValValTyrHisGly 241
QY 406 CATTTC-----TCCACCAAGAACAACCAATGGACTATGATATTTGCCCTTTTG 450
Db 242 GlyTyrLeuProPheArgAspProAsnSerGluGlnAsnSerAsnAspIleAlaLeuVal 261
QY 451 AAGATGGCTGGAGCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCTCCAGAGCTG 510
Db 262 HisLeuSerSerProLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 281
QY 511 CGGAGCAATTTGAGGCTGTTTATTGTACAACTCCAGGCTGGGGCCGCTTAACGTAA 570
Db 282 GlyGlnAlaLeuValaspGlyLysIleCysThrValThrGlyTrpGlyAsnThrGlnTyr 301
QY 571 GGTGGCTCTCTCAAGTCTTTCAGGAAGTGAATCTGCCTATTGTTGACCTGGGAAGAG 630
Db 302 TyrGlyGlnGlnAlaGlyValLeuGlnGlnAlaArgValProIleIleSerAsnAspVal 321
QY 631 TGTGTGCAGCTCTGTTAAACATAAGAGGCCCATCATGCTGGGAGAGACCTTTCTTTCAC 690
Db 322 CysAsnGlyAlaAspPhe---TyrGlyAsnGlnIleLysProLysMetPhe---CysAla 339
QY 691 GGTTCCTCTGATGAGGAGGAGCATGTCAGGAGATTCAGGAGGTTCACTCATGTGC 750
Db 340 GlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSerGlyGlyProPheValCys 359
QY 751 CGG-----AATAAGAAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
Db 360 GluAspSerIleSerArgThrProArgTyrArgLeuCysGlyIleValSerTrpGlyThr 379
QY 802 GGCTGTGCTGCTGAGCTGGAGAAACAATGTGAGGAAAAGTATCAAGATCCCTCGGATC 861
Db 380 GlyCysAlaLeuAlaGlnLys-----ProGlyVal 399
QY 862 TTCACAGACATTAAGTAAAGTCTTCTCTGATCCAGACACATCCAACT 912
Db 390 TyrThrLysValSerAspPheArgGluTrpIlePheGlnAlaIleLysThr 406

RESULT 13
HEPS_RAT
ID HEPS_RAT STANDARD; PRT; 416 AA.
AC Q05511;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease hepsin (EC 3.4.21.-).
GN HPN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93305733; PubMed=8318546;
RA Farley D., Reymond F., Nick H.;
RT "Cloning and sequence analysis of rat hepsin, a cell surface serine
RT proteinase";
RL Biochim. Biophys. Acta 1173:350-352(1993).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE
CC OF CELL MORPHOLOGY.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)

CC or send an email to license@isb-sib.ch.
CC EMBL; X70900; CAA50256.1; -.
DR PIR; S32013; S32013.
DR PIR; S33777; S33777.
DR HSP: P00763; IDPO.
DR MEROPS; S01.224; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR01254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Transmembrane; Signal-anchor.
FT CHAIN 1 161 NON-CATALYTIC CHAIN (POTENTIAL).
FT CHAIN 162 416 CATALYTIC CHAIN (POTENTIAL).
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 44 416 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 162 416 SERINE PROTEASE.
FT ACT_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 256 256 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 352 352 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 152 276 INTERCHAIN (BY SIMILARITY).
FT DISULFID 187 203 BY SIMILARITY.
FT DISULFID 321 337 BY SIMILARITY.
FT DISULFID 348 380 BY SIMILARITY.
FT CARBOHYD 111 111 N-LINKED (GLCNAC... (POTENTIAL).
SQ SEQUENCE 416 AA; 44926 MW; E5A9F8FA9550E180 CRC64;

Alignment Scores:
Pred. No.: 4,68e-33 Length: 416
Score: 438.00 Matches: 99
Percent Similarity: 49.47% Conservative: 41
Best Local Similarity: 34.98% Mismatches: 99
Query Match: 26.26% Indels: 44
DB: 1 Gaps: 10

US-09-735-713a-1 (1-921) x HEPS_RAT (1-416)

QY 109 TGTGGCAGAGTCTGTTAAAGTACAGCCTTGAATTTATTTAACTTTTCAGTCGCATT 168
Db 152 CysGlyArgArgLysLeuProVal-----AspArgIle 162
QY 169 CTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGG 228
Db 163 ValGlyGlyGlnAspSerSerLeuGlyArgTyrProTyrGlnValSerLeuArgTyrAsp 182
QY 229 CAGAAGCATTTTGGGAGGAGCATCTCTCACCACAGTGGGTGATCAGGGGGCTCAC 288
Db 183 GlyThrHisLeuCysGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAlaAlaHis 202
QY 289 TCATTGCAACACAGAAC---ATTGTGCTCTACTTTCAATTTACTCTCTGGAGAGTATGAC 345
Db 203 CysPheProGluArgAsnArgValLeuSerArgTyrArgValPheAlaGlyAla----- 220
QY 346 TTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAACTGTCAATCATATCATCA 405
Db 221 ValAlaArgThrSerProHisAlaValGlnLeuGlyValGlnAlaValIleTyrHisGly 240
QY 406 CATTTC-----TCCACCAAGAACAACCAATGGACTATGATATTTGCCCTTTTG 450
Db 241 GlyTyrLeuProPheArgAspProThrIleAspGluAsnSerAsnAspIleAlaLeuVal 260
QY 451 AAGATGGCTGGAGCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCTCCAGAGCTG 510
Db 261 HisLeuSerSerSerLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 280

```
Qy 511 CGGAGCAATTGAGCGCTGTTTATTGTTACAACTGAGCGTGGCGCTTAACCTGAA 570
Db 281 GlyGlnAlaLeuValAspGlyValCysThrValThrGlyTrpGlyAsnThrClnPhe 300
Qy 571 GGTGGCGTCTCTCACAAGCTTTCAGGAAGTGAATCTGCTATTTTGAACCTGGGAAGAG 630
Db 301 TyrGlyGlnGlnAlaValValLeuGlnGlnAlaArgValProIleIleSerAsnGluVal 320
Qy 631 TGTGTGGCAGCTCTCTTAACACTAAGAGGCC-----ATCAGTGGG 672
Db 321 Cys-----AsnSerProAspPheTyrGlyAsnGlnIleLysPro 333
Qy 673 AAGACCTTTCTTGTGCACAGTTTCTGATGAGGAGAGACGCGATGTCAGGAGATTCATCA 732
Db 334 LysMetPhe---CysAlaGlyTyrProGluGlyIleAspAlaCysGlnGlyAspSer 352
Qy 733 GGAGTTTCACATCAGTGGCGGAATAAG-----AAAGGGCGCTGACCTCTGGCTGGT 783
Db 353 GlyGlyHisPheValCysGluAspArgIleSerGlyThrSerArgTrpArgLeuCysGly 372
Qy 784 GTGACTTCTCTGGGTTTGGCTGGTGGTGGAGGAGGAGAAACAATGTGAGGAAAGTGTAT 843
Db 373 IleValSerTrpGlyThrGlyCysAlaLeu-----AlaArgLys----- 385
Qy 844 CAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTCTTCTGATGATCCAGCAACAC 903
Db 386 -----ProGlyValTyrThrLysValIleAspPheArgGluTrpIlePheGlnAla 402
Qy 904 ATCCAACT 912
Db 403 IleLysThr 405

RESULT 14
TMS5_HUMAN STANDARD; PRT; 457 AA.
ID Q9H3S3;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).
GN TMRSS5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=11741986;
RA Yamaguchi N., Okui A., Yamada T., Nakazato H., Mitsui S.;
RT "Spinesin/TMRSS5, a novel transmembrane serine protease, cloned from
RT human spinal cord."
RL J. Biol. Chem. 277:6806-6812(2002).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Brain-specific. Predominantly expressed in
CC neurons, in their axons, and at the synapses of motoneurons in the
CC spinal cord.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB028140; BAB20375.1; -
DR HSSP; P00763; LDPO.
DR Genew; HGNC:14908; TMRSS5.
DR MIM; 606751; -
DR MEROPS; S01.313; -
```

```
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; SR; 1.
DR SMART; SM00202; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS00287; SRCR_2; FALSE_NEG.
DR Hydrolase; Serine protease; Transmembrane; Signal-anchor;
KW Glycoprotein.
FT DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 50 70 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 71 457 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 112 207 SRCR.
FT DOMAIN 218 457 SERINE PROTEASE.
FT ACT_SITE 258 258 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 308 308 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 405 405 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 217 218 CLEAVAGE (POTENTIAL).
FT DISULFID 135 196 BY SIMILARITY.
FT DISULFID 148 206 BY SIMILARITY.
FT DISULFID 209 328 BY SIMILARITY.
FT DISULFID 243 259 BY SIMILARITY.
FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 401 429 BY SIMILARITY.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 457 AA; 49574 MW; 64406AB4985A2651 CRC64;

Alignment Scores:
Pred. No.: 4.78e-33 Length: 457
Score: 438.00 Matches: 93
Percent Similarity: 53.91% Conservative: 45
Best Local Similarity: 36.31% Mismatches: 94
Query Match: 26.26% Indels: 24
DB: 1 Gaps: 8

US-09-735-713A-1 (1-921) x TMS5_HUMAN (1-457)
Qy 160 AGTCGCAATCTTGGAGGAAGCCAGTGGAGAGGGTCTCTATCCCTGCAGGATATCTGT 219
Db 216 SerArgIleValGlyGlyGlnSerValAlaProGlyArgTrpProTrpGlnAlaSerVal 235
Qy 220 AAACAAGGCGAGACATATTTGTGGAGGAAGCATCGTCTCACCACAGTGGGTGATCAGG 279
Db 236 AlaLeuGlyPheArgHisThrCysGlyGlySerValLeuAlaProArgTrpValValThr 255
Qy 280 CGGCGTCACTGCATTCGAAACAGAACATTT-----GTGCTCTACTTTGAATGTTACTGCT 333
Db 256 AlaAlaHisCysMetHisSerPheArgLeuAlaArgLeuSerSerTrpArgValHisAla 275
Qy 334 GGAGAGTATGACTTAAGCCAGACACGCCAGGAGAGCAAACTCTCATTATTTGAAGTGTCT 393
Db 276 GlyLeuValSerHisSerAlaValArgProHisGlnGlyAlaLeu---ValGluArgIle 294
Qy 394 ATCATACATCCACATTTCTCCACCAGAAACCAATGGACTATGATATGCCCTTTTGAAG 453
Db 295 IleProHisProLeuTyrSerAlaGlnAsn---HisAspTyrAspValAlaLeuLeuArg 313
Qy 454 ATGGCTGAGCGCTTCCAATTTGGGCCCATCTTTGTGGGCCCATATGTTCTCCAGAGCTGGG 513
Db 314 LeuGlnThrAlaLeuAsnPheSerAspThrValGlyAlaValCysLeuProAlaLysGlu 333
Qy 514 GAGCAATTTGAGCGTGGTGTATTTTATTTTACAACATGCAGGCTGGGGCGCTTAACTGAAGGT 573
```

```

Db 334 GlnHisPheProLysGlySerArgCysTrpValSerGlyTrpGlyHisThrHisProSer 353
QY 574 GGCGTCTCTC---TCACAGCTTCGACGAGTGAATCGCTATTTGACCTGGGAAGAG 630
Db 354 HisThrTyrSerSerAspMetLeuGlnAspThrValValProLeuPheSerThrGlnLeu 373
QY 631 TGTGTGGCAGCTGTAAACACTAAAGAGGCCCATCAGTGGG-----AAGACCTTT 681
Db 374 CysAsnSerSerCysVal-----TyrSerGlyAlaLeuThrProArgMet 388
QY 682 CTTTGGCAGCTTCTCTGATGAGGAGACGCATGTCAGGAGATCAGGAGCTTCA 741
Db 389 LeuCysAlaGlyTyrLeuAspGlyArgAlaAspAlaCysGlnGlyAspSerGlyPro 408
QY 742 CTCATGTGCGGGAATAAGAAAGGGCGCTGACTCTGGCTGGTGTGACTCTCCTGGGTTG 801
Db 409 LeuValCysProAsp---GlyAspThrTrpArgLeuValGlyValSerTrpGlyArg 427
QY 802 GGCTGTGTGCGAGCTGGGAGAACAAATGTGAGGAAAGTGAAGGATCCCTGGGATC 861
Db 428 AlaCys-----AlaGluProAsnHisProGlyVal 437
QY 862 TTCACAGACATTAGTAAGTGCCTTCTCGATCCAGCAACACATCCAA 909
Db 438 TyrAlaLysValAlaGluPheLeuAspTrpPileHisAspThrAlaGln 453

RESULT 15
FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN Fl2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94242782; PubMed=8186251;
RA Shibuya Y., Sempa U., Okabe H., Kambara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor XII): comparison with human and guinea pig molecules.";
RL Biochim. Biophys. Acta 1206:63-70(1994).
RN [2]
RP SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77182112; PubMed=861210;
RA Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman factor).";
RL Biochemistry 16:2270-2278(1977).
CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIa and factor XI to form factor Xla.
CC -1- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XI FIRST TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/KALLIKREIN CLEAVAGE SITE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.

```

```

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; S70164; AAB30804.2; -.
CC HSP; P00763; IDPO.
CC MEROPS; S01.211; -.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR000083; Fibrinctn1.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC ProDom; PD000395; Kringle; 1.
CC ProDom; PD000995; FN_Type_II; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS0240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT SIGNAL 1 9
FT CHAIN 10 349
FT CHAIN 350 593
FT DOMAIN 37 78
FT DOMAIN 84 121
FT DOMAIN 123 163
FT DOMAIN 164 200
FT DOMAIN 207 287
FT DOMAIN 297 333
FT DOMAIN 350 593
FT ACT_SITE 389
FT ACT_SITE 438
FT ACT_SITE 541
FT DISULFID 88 100
FT DISULFID 94 109
FT DISULFID 111 120
FT DISULFID 125 153
FT DISULFID 151 160
FT DISULFID 168 179
FT DISULFID 173 188
FT DISULFID 190 199
FT DISULFID 207 287
FT DISULFID 230 269
FT DISULFID 258 282
FT DISULFID 336 463
FT DISULFID 374 390
FT DISULFID 382 452
FT DISULFID 413 416
FT DISULFID 479 547
FT DISULFID 510 526
FT DISULFID 537 568

```

Alignment Scores:

Pred. No.:	1.48e-32	Length:	593
Score:	433.00	Matches:	101
Percent Similarity:	52.14%	Conservative:	45
Best Local Similarity:	36.07%	Mismatches:	94
Query Match:	25.96%	Indels:	40
DB:	1	Gaps:	12

	QY	109	TGTGGCGAGAGTCTGGTTTAGGTACAGCCCTTGGAATATTATTTAACATATTTTCAGTCGCATT	168
	DB	336	CysGlyGlnArgLeuArgLys-----TrpLeuSerSerLeuAsnArgVal	350
	QY	169	CTTGGAGGAACCAAGTGGAGAAGGCTTCTATCCTCGCAGGTACTCTGAAACAAGG	228
	DB	351	ValGlyGlyLeuValAlaIeuProGlyAlaHisProTyrIleAlaAlaIeuTyrTrpAsp	370
	QY	229	CAGAAGCATATTTGTGGAGGAAGCATCGTCTCCACAGTGGGTGTATCACGGCGGCTCAC	288
	DB	371	Gln--HisPheCysAlaGlySerLeuIleAlaProCysTrpValLeuThrAlaAlaHis	389
	QY	289	TGCATTGCCAACAACAACATTGTGTCTACTTTCGAATGTTCCTCGGAGAGTATGACTTA	348
	DB	390	CysLeuGlnAsnArgProAlaProLysGluLeuThrValValLeuGlyGlnAspArgHis	409
	QY	349	AGCCAGACAGACCAGGAGAGCAAACTCTCATCTATTGAAACTGTCTATCATATCCACAT	408
	DB	410	AsnGlnSerCysGluGlnCysGlnThrLeuAlaValArgAspTyrArgLeuHisGluAla	429
	QY	409	TTCTCCACCACCAACCAATGGACTAT-----GATATTGCCCTTTTGAAGATGGCTGA	462
	DB	430	PheSer-----ProIleThrTyrGlnHisAspLeuAlaLeuValArgLeuGlnGlu	446
	QY	463	GCCTTCCTCAA-----TTTGTGGGGCCCCATATGTCTCTCCAGAG	507
	DB	447	SerAlaAspGlyCysCysAlaHisProSerProPheValGlnProValCysIeuProSer	466
	QY	508	-----CTGGGGGAGCAATTTAGCGCTGGTTTTATTGTACACTGCAGGCTGGGGCCGC	561
	DB	467	ThrAlaAlaArgProAlaGluSerGluAlaAlaValCysGluValAlaGlyTrpGlyHis	486
	QY	562	TTAACTGAGGTGCGCTCTCTCACAACTCTCCAGGAAGTGNATCTGCCTATTTTGACC	621
	DB	487	GlnPheGluGlyGlyGluTyrSerSerPheLeuGlnGluAlaGlnValProLeuIleAsp	506
	QY	622	TGGGAAGAGTGTCTGGCAGCTCTCTTAACACTAAAGAGGCCATAGTGGGAAGACTTT	681
	DB	507	ProGlnArgCysSerAlaPro-----AspValHisGlyAlaAlaPhe	520
	QY	682	-----CTTTGCACAGGTTTTCTGTATGGAGGGAGACGCATGTCTCAGGGAGAT	729
	DB	521	ThrGlnGlyMetLeuCysAlaGlyPheLeuGluGlyGlyThrAspAlaCysGlnGlyAsp	540
	QY	730	TCAGGAGGTTCACTCATGTGCCGGAATAAG-----AAAGGGCCCTGGACTCTGGCTGGT	783
	DB	541	SerGlyGlyProLeuValCysGluAspGluThrProGluArgGlnLeuIleLeuArgGly	560
	QY	784	GTCACTTCTCTGGGTTTGGCGCTGTGTCTCGAGGCTGGAGAAACAATCTGAGGAAAAGT	843
	DB	561	IleValSerTrpGlySerGlyCysGly-----AsnArgLeuLys-----	573
	QY	844	CAAGGATCCCTCGGATCTTCACAGACATTTAGTAAGTGTCTTCCTGGGATCCACGAAC	903
	DB	574	-----ProGlyValTyrAspValAlaAsnTyrLeuAlaTrpIleArgGluHis	590

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 25, 2003, 06:51:58 ; Search time 90.5 seconds

(without alignments)

4193.799 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 1668

Sequence: 1 atgagtctcaaatgcttat.....acatccaaactggtactaa 921

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09735713/runat_18032003_124354_23129/app_query.fasta_1.1095
-DB=SPTRMBL_21 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713 @CGN_1_1_57 @runat_18032003_124354_23129 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	690.5	41.4	1004	13	P79953	P79953 xenopus lae

2	657	39.4	974	13	Q90WD8	Q90wd8 bufo japoni
3	626	37.5	1524	13	Q91674	Q91674 xenopus lae
4	496	29.7	277	5	Q96899	Q96899 scolopendra
5	486	29.1	787	5	Q9VEY6	Q9vey6 drosophila
6	480.5	28.8	799	11	Q9DBI0	Q9dbi0 mus musculus
7	475.5	28.5	335	11	Q8VIF2	Q8vif2 mus musculus
8	473.5	28.4	624	11	Q9DAT3	Q9dat3 mus musculus
9	467.5	28.0	624	11	Q91V47	Q91v47 mus musculus
10	462.5	27.7	467	5	Q967X8	Q967x8 panulirus a
11	462	27.7	453	11	Q8VDE0	Q8vde0 mus musculus
12	458	27.5	638	11	Q8R0P5	Q8rup5 mus musculus
13	450	27.0	339	11	Q99L44	Q99l44 mus musculus
14	446	26.7	643	6	Q97506	Q97506 sus scrofa
15	441.5	26.5	433	13	Q90VK1	Q90vk1 brachydanio
16	436.5	26.2	317	13	Q9DGR3	Q9dgr3 xenopus lae
17	430	25.8	767	13	Q9DGR2	Q9dgr2 xenopus lae
18	429.5	25.7	310	11	Q91XC4	Q91xc4 mus musculus
19	428.5	25.7	581	5	Q960I5	Q960i5 drosophila
20	428.5	25.7	1047	5	Q9VZH2	Q9vzh2 drosophila
21	427.5	25.6	387	5	Q9XY57	Q9xy57 ctenecephal
22	426	25.5	624	6	Q95ME7	Q95me7 oryctolagus
23	424.5	25.4	310	11	Q9QYZ9	Q9qyz9 mus musculus
24	423.5	25.4	616	6	Q97507	Q97507 sus scrofa
25	422.5	25.3	267	5	Q9BK47	Q9bk47 luidia foli
26	422.5	25.3	456	6	Q9TTR0	Q9tttr0 canis famil
27	422	25.3	314	5	Q9VR15	Q9vr15 drosophila
28	421	25.2	273	6	Q9XSM1	Q9xsm1 ovis aries
29	421	25.2	855	11	Q9JJ17	Q9jj17 rattus norv
30	420.5	25.2	1047	5	Q24019	Q24019 drosophila
31	420	25.2	321	4	Q96RZ8	Q96rz8 homo sapien
32	419	25.1	256	5	Q9XYV1	Q9xyv1 rhyzopertha
33	418.5	25.1	325	5	O15944	O15944 sarcophaga
34	418.5	25.1	329	6	Q9GL10	Q9gl10 ovis aries
35	417	25.0	284	11	Q9DP78	Q9dp78 mus musculus
36	417	25.0	264	11	Q9ER05	Q9er05 mus musculus
37	414.5	24.9	300	4	Q96EF3	Q96ef3 homo sapien
38	414.5	24.9	573	5	Q9V516	Q9v516 drosophila
39	414	24.8	254	5	Q8T637	Q8t637 aedes aegyp
40	414	24.8	264	11	Q9EQZ8	Q9eqz8 rattus norv
41	413.5	24.8	288	5	O46151	O46151 pacifastacu
42	413	24.8	263	11	Q9DC86	Q9dc86 mus musculus
43	413	24.8	405	4	Q96E86	Q96e86 homo sapien
44	412	24.7	263	11	Q9CR35	Q9cr35 mus musculus
45	409.5	24.6	468	5	Q9U0G3	Q9u0g3 pacifastacu

ALIGNMENTS

RESULT 1

P79953	ID	P79953	PRELIMINARY;	PRT;	1004 AA.
AC	P79953;				
DT	01-MAY-1997	(TREMBLrel. 03, Created)			
DT	01-MAY-1997	(TREMBLrel. 03, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	Oviductin.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;				
OC	Xenopodinae; Xenopus.				
OX	NCBI_TaxID=8355;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99184825; Pubmed=10084976;				
RA	Lindsay L.L., Wieduwilt M.J., Hedrick J.L.;				
RT	"Oviductin, the Xenopus laevis oviductal protease that processes egg				
RT	envelope glycoprotein gp43, increases sperm binding to envelopes, and				
RT	is translated as part of an unusual mosaic protein composed of two				
RT	protease and several CUB domains."				
RL	Biol. Reprod. 60:989-995(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Yang J.C., Lindsay L.L., Hedrick J.L.;				

Db	197	AspLeuSerSerCysLeuHisValMetSerAlaLeuLysGlyThrValValSerSerTyr	216
QY	679	TTTCTTTGCACAGGTTTTCTGTGATGGAGGAGACGACATGTCTCAGGGAGATTCTCAGGAGGT	738
Db	217	IleValCysAlaGlyPheProGluGlyClyLysAspAlaCysGlnGlyAspSerGlyGly	236
QY	739	TCACATCATGTCCGGGATATAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTCTCTCGGGT	798
Db	237	ProLeuLeuCysGlnArgHisGlySerTyrValLeuHisGlyLeuThrSerTyrGly	256
QY	799	TGCGCTGTCTCAGGCTGCAGAGAACAAATGTG-----AGGAAAGTGTATCAAGATCC	852
Db	257	MetGlyCysGlyArgSerTyrPlyAsnAsnValPheLeuProHisAsnArgLysGlySer	276
QY	853	CCTGGGATCTTCACAGACATAGTAAAGTGTCTCTCTGGATCCACGAACATCCAAACT	912
Db	277	ProGlyIlePheThrAspIleGlnLysLeuLeuGlyTyrValSerSerGlnLeuAsnThr	296
RESULT	2		
Q90WD8			
ID	Q90WD8	PRELIMINARY;	PRT; 974 AA.
AC	Q90WD8;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Oviductin.		
OS	Bufo japonicus (Japanese toad).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae; Bufo		
OX	NCBI_TaxID=8387;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=OVIDUCTAL PARS RECTA;		
RA	Hiyoshi M., Takamune K., Mita K., Kubo H., Sugimoto Y., Katagiri C.;		
RT	Oviductin, the oviductal protease that mediates gamete interaction b		
RT	affecting the vitelline envelope in Bufo japonicus: Its molecular		
RT	cloning and analyses of expression and post-translational		
RT	activation.;		
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB070367; BAB63372.1; -		
DR	MEROPS; S01.240; -		
DR	InterPro; IPR000859; CUB_domain.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00431; CUB; 2.		
DR	Pfam; PF00089; trypsin; 2.		
DR	PROSITE; PS01180; CUB; 2.		
DR	PROSITE; PS50240; TRYPSIN_DOM; 2.		
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.		
DR	PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.		
KW	Hydrolase; Serine protease.		
SQ	SEQUENCE 974 AA; 107647 MW; F19705A470465553 CRC64;		
Alignment Scores:			
Pred. No.:	3.65e-59	Length:	974
Score:	657.00	Matches:	116
Percent Similarity:	62.12%	Conservative:	66
Best Local Similarity:	39.59%	Mismatches:	103
Query Match:	39.39%	Indels:	8
DB:	13	Gaps:	3
US-09-735-713A-1 (1-921) x Q90WD8 (1-974)			
QY	34	CTGATTTTACTACTAGGATAGTCTTTTGTACRAGGTAAATCTGCARCTCTTTCGCTC	93
Db	12	LeuThrValMetIleGlyVal-----GlyArgGlyValThrAspSerPro	26
QY	94	CCCAAGACTCCAGTGTGTGGCAGAGCTGTGTTAAGGTACAGCTTGGAAATATTAAAC	153
Db	27	GlyArgValSerArgCysGlyGluArgProAlaAlaAsnThrSerValSerTyr---Gly	45
QY	154	ATTTCAGTCGCATCTTGTGGAGAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTA	213

```
Db 46 LeuLeuSerArgIleValGlyGlyThrSerAlaValLysGlyGluSerProTrrpMetVal 65
Qy 214 TCTCTGAACAAAGCAGACATATTGTGGAGAGCATCGCTCACCACAGCTGGTG 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 SerLeuLysArgAspGlyLysHisPheCysGlyGlyThrIleLeuSerAspLysTyrVal 85
Qy 274 ATCAGCGGCTCACTGCATGTGCAACAGACAACATTTGTCTACTTTGAATGTTACTGCT 333
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 86 LeuThrAlaAlaHisCysValLeuGluLysAsnPheGluPheGlnValSerValSerIle 105
Qy 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCATTATGAACATGTC 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 GlyAspHisAspPheAlaValTyrGluArgSerGluGlnArgPheAlaIleLysSerVal 125
Qy 394 ATCATACATCACATTCTCCACCAAGAAACCAATGACTATGATATTGCGCTTTTGAAG 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 PheLysHisProAsnPheLysProSerArgProPheAsnTyrAspLeuAlaIleLeuGlu 145
Qy 454 ATGGCTGGAGCCTTCCAATTGGCCACTTTGTGGGGCCCATATGCTTCTCCAGAGCTCGG 513
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 146 LeuValGluSerIleThrPheAspLysAspIleGlnProAlaCysLeuProSerProAsp 165
Qy 514 GAGCAATTGAGGCTGGTTTATTTGTACAACTCCAGGCTGGGGCCGCTTAACGTGAAGGT 573
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 AspValPheProThrGlyThrLeuCysMetAlaLeuGlyTrpGlyArgLeuGlnGluAsn 185
Qy 574 GCGCTCTCTCACAAAGCTTGCAGAAAGTGAATCTGCCTATTGACCTGGGAAGAGTGT 633
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 GlyArgLeuProSerSerLeuGlnLysValValLeuProLeuIleGluTyrArgArgCys 205
Qy 634 GTGGCAGCTCTGTAAACACTAAAGAGGCCCATCATCAGTGGGAGACCTTTCTTTCACAGGT 693
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 206 LeuSerIleMetGluThrValAspArgLeuAlaPheGluThrValValCysAlaGly 225
Qy 694 TTTCTCTATGAGGAGGAGCAGCCATGTCAGGAGATTCAGGAGTTCACTCATCTGCGCG 753
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 PheProGluGlyGlyLysAspAlaCysGlnGlyAspSerGlyGlyProPheLeuCysGln 245
Qy 754 AATAAGAAAGGGCTGGACTCTGGCTGGTGTGACTTCTCTGGGCTTGGGCTGGTGGCGA 813
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 ArgSerGlnGlyArgTrpValLeuValGlyValThrSerTrpGlyLeuGlyCysAlaArg 265
Qy 814 GGCTGGAGAAACAATGTGAGGAAA-----AGTGATCAAGGATCCCTGGGATCTTCACA 867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 266 LysTrpAlaAspAsnIleLeuAspProValGluSerLysGlySerProGlyValPheThr 285
Qy 868 GACATTAGTAAGAGCTTCTCTGATCCAGCAACACATC 906
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 AspIleGlnArgLeuLeuAsnTrpLeuSerGluAsnLeu 298

RESULT 3
Qy1674 ID Q91674 PRELIMINARY; PRT; 1524 AA.
AC Q91674;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polyprotein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99432219; PubMed=10500163;
RT Lindsay L.L., Yang J.C., Hedrick J.L.;
RT "Ovochymase, a Xenopus laevis egg extracellular protease, is
RT translated as part of an unusual polyprotease.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:11253-11258(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX Yang J.C., Lindsay L.L., Hedrick J.L.;
```

```
RT "cDNA Cloning of Ovochymase, a Chymotrypsin-like Protease Released
RT From Xenopus laevis Eggs at Fertilization.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 CUB DOMAINS
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U81290; AAC24717.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.022; -.
DR MEROPS; S01.245; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001254; Ser_protease_try.
DR Pfam; PF00431; CUB; 5.
DR Pfam; PF00089; trypsin; 3.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; Tryp_SPC; 3.
DR SMART; SM00020; CUB; 4.
DR PROSITE; PS01180; CUB; 5.
DR PROSITE; PS00134; TRYPSIN_DOM; 3.
DR PROSITE; PS00135; TRYPSIN_SER; 3.
DR Hydrolase; Serine protease.
KW CHAIN 57 308 SERINE PROTEASE.
FT CHAIN 584 817 SERINE PROTEASE.
FT CHAIN 1295 1524 OVOCHYMASE.
SQ SEQUENCE 1524 AA; 167566 MW; 32EFE42128F37268 CRC64; )

Alignment Scores:
Pred. No.: 6,75e-56 Length: 1524
Score: 626.00 Matches: 128
Percent Similarity: 57.89% Conservative: 37
Best Local Similarity: 44.91% Mismatches: 110
Query Match: 37.53% Indels: 10
DB: 13 Gaps: 4

US-09-735-713a-1 (1-921) x Q91674 (1-1524)
Qy 70 GGTAAATCTGCARCTCTTCGCTCCCAAGAGCTCCAGTGTGGCAGAGTCTGTTAAG 129
    |||:||||| ||| ||| ||| |||
Db 32 GlyIuThrAlaGluLeuLysCysGlyThrArgProGluIleGlyAsp----- 47
Qy 130 GTACAGCCTTGGAAATTATTTAACATTTTCAGTCGCATTTCTTGGAGGAAGCAAGTGGAG 189
    ::||| ||||| ||||| ||||| |||||
Db 48 --GluProAspLeuGluPheThr-----SerArgIleValGlyGlyAspAlaAla 64
Qy 190 AAGGTTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGCAAGCATATTTGTGGAGGA 249
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 ValGlyGlyGlnProTrpThrValSerLeuLysLeuAsnGluArgHisIleCysGlyGly 84
Qy 250 AGCATCGTCTCACACAGTGGTGCATCAGCGCGCTCACTGCATT-----GCAAAACAGA 303
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 SerIleValArgLysAspMetValThrAlaAlaHisCysValTyrProValThrGlu 104
Qy 304 AACATTGTGTCTACTTTGAATTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCA 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 IleLysValSerHisMetThrValIleValGlyGlyTrpAspGlnGlnValMetAspSer 124
Qy 364 GGAGAGCAAACTCTCACTATTGAACACTGTCATCATCATCATCTCCACCAAGAAA 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 GlnGluGlnSerIleProValSerHisIleGluProHisProAsnTyrArgGlyAspGly 144
Qy 424 CCAATGACTATGATATTGCGCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCACTTT 483
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 AsnMetGlyTyrAspIleAlaLeuValPheLeuSerLysProIleIlePheGlySerGln 164
Qy 484 GTGGGGCCCATATGCTTCTCCAGAGCTCGGGAGCAATTTAGCGTGGTGTATTATTGTACA 543
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 165 ValGlnProIleCysLeuProGlnValGlyGlyLysIleGluAlaGlyThrLeuCysVal 184
Qy 544 ACTCAGGCTGGGGCCGCTTAACCTGAAGGTGGCGCTCTCTCACAGAGTCTTCACGAAGTG 603
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 SerSerGlyTrpGlyArgLeuGluGluAsnGlyAspLeuSerProValGluGlnVal 204
```

```

QY 604 AATCTGCTATTTTACCTGGGAGAGTGTGTGCGAGCTCTGTAAACACTAAAGAGGCC 663
    |||||:||||:
Db 205 LysLeuProValValAspAsnGlyThrCysHisAlaValLeuGluProIleGlyHisPro 224
QY 664 ATCAGTGGGAGACCTTTCTTCCACAGGTTTCTCGATGGAGGAGAGACGATGTCAG 723
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 225 ValLeuAspAspThrMetLeuCysAlaGlyPheProGluGlyGlyMetAspAlaCysGln 244
QY 724 GGAGATTACAGAGGTTCACTCATGTGCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGT 763
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 245 GlyAspSerGlyGlyPropheValCysArgArgSerGlyValThrPheLeuAlaGly 264
QY 784 GTGACTTCTGGGTTTGGGCTGTGTGCGAGCTGG---AGAAACAATGTGAGAAAGT 840
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 265 CysValSerTrpGlyLeuGlyCysGlyArgSerTrpGlyAlaLysGlnIleIleArgSer 284
QY 841 GATCAAGGATCCCTGGGCTTTCACAGACATAGTAAGAGTGTTCCTGGATCCACGAA 900
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 285 GlnSerGlySerProAlaIlePheSerArgValSerSerValLeuAspPheLeuArgPro 304
QY 901 CACATCCAAACTGGT 915
Db 305 ProLysLeuThrcly 309

RESULT 4
ID O96899 PRELIMINARY; PRT; 277 AA.
AC O96899;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Plasminogen activator SPA.
OS Scolopendra subspiniipes.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Scolopendra.
OX NCBI_TaxID=55038;
RN [1]
RP SEQUENCE FROM N.A.
RA Sohn Y.-D., You W.K., Kim K.Y., Chung K.H., Park D.-H.;
RT "Cloning of a cDNA for a novel plasminogen activator from a Korean
RT centipede Scolopendra."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U79521; AAD00320.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.122; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 277 AA; 30172 MW; EC62F80C8CB246C CRC64;

Alignment Scores:
Pred. No.: 1.37e-42 Length: 277
Score: 496.00 Matches: 112
Percent Similarity: 53.74% Conservative: 46
Best Local Similarity: 38.10% Mismatches: 98
Query Match: 29.74% Indels: 38
DB: 5 Gaps: 12

US-09-735-713A-1 (1-921) x O96899 (1-277)
QY 28 AACAACTGATTTTACTACTAGGAATGCTTTTGTGAACRAGGTAAATCTGCARCTCT 87
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2 AsnSerPheThrIleLeuIle---ValThrTyrPhe----- 12

```

```

QY 88 TCGCTCCCCAAAGCTCCCACTGTTGTGGCAGAGTCTGTTAAGGTACAGCCTTGGAAATTAT 147
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 13 SerLeuAlaPheGlySerArgCysGly-----IleLysAsnGlyPro-----Met 27
QY 148 TTTAACTATTTTCAGTCGCATTTCTTGGAGGAGCCAAAGTGGAGAGGTTCCCTATCCCTGG 207
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 LeuAspGluPheAsnArgIleValGlyGlyGluAlaAlaGluProGlyGluPheProTrp 47
QY 208 CAGGTATCTCTGAAA-----CAAAGGCAGAACATATTGTGGAGGAACATC 255
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 48 GlnIleSerLeuGlnValValSerTrpTyrGlySerTyrHisTyrCysGlyGlySerIle 67
QY 256 GTCTACCAACAGTGGGTGATCATCGCGCGCTCACTGCATTGCCAACAGAAACATTTGTCT 315
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 68 LeuAspGluSerTrpValThrAlaAlaHisCysValGluGlyMetAsn---ProSer 86
QY 316 ACTTTGAATGTTACTCTGAGAGTATGACTTAAGCCACACAGACACCCAGAGAGAAACT 375
    |||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 87 AspLeuArgIleLeuAlaGlyGluHisAsnPheLysLysGluAspGlyThrGluGlnTrp 106
QY 376 CTCACATATTGAAACTGTCATCATACATCCACATTTCTCCACCACCAAGAAACCAATGGACTAT 435
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 107 GlnAspValIleAspIleIleMetHisLysAspTyr---ValTyrSerThrLeuGluAsn 125
QY 436 GATATTGCCCTTTTGAAGATGGCTGGAGCTTCCAAATTT---GGCCACTTTGTGGGGCCC 492
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 126 AspIleAlaLeuLeuLysLeuAlaGluProLeuAspLeuThrProThrAlaValGlySer 145
QY 493 ATATGCTTTCCAGAGCTCGGGAGCAATTTGAGGCTGTTTATTTTGTACAACTGCAGGC 552
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 146 IleCysLeuProSerGlnAsnAsnGlnGluPheSerGlyHis---CysIleValThrGly 164
QY 553 TGGGGCGCTTAACCTAAGTGGCTCTCACAACTCTTGCGAGGAAGTGAATCTGCCT 612
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 165 TrpGlySerValArgGluGlyGlyAsnSerProAsnIleLeuGlnLysValSerValPro 184
QY 613 ATTTTGACCTGGGAGAGTGTGTGGCAGCTCTGTTAACACATAAGAGGCCCATCAGTGGG 672
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 185 LeuMetThrAspGluGluCysSerGluTyrTyrAsnIleVal----- 198
QY 673 AGACCTTTCTTTCACAGCTTTTCTGATGGAGGAGAGACGCATCTCAGGAGATTCATCA 732
    |||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 199 AspThrMetLeuCysAlaGlyTyrAlaGluGlyGlyLysAspAlaCysGlnGlyAspSer 218
QY 733 GGAGGTTCATCATGTGCGCGGAATAAGAAAGGGCCCTGGACTCTGCTGGTGTGACTTCC 792
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 219 GlyGlyProLeuValCysProAsnGlyAspGlyThrTyrSerLeuAlaGlyIleValSer 238
QY 793 TGGGGTTTGGCTGTGTCAGGCTGGAGAAACAATGTGAGGAAAAAGTATCAAGGATCC 852
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 239 TrpGlyIleGlyCysAlaGlnPro---ArgAsn----- 248
QY 853 CTGGGATCTTCACAGACATATTAGTAAAGTCTTCTTCTGGATC 894
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 249 ProGlyValTyrThrGlnValSerLysPheLeuAspTrpIle 262

RESULT 5
Q9VEY6
ID Q9VEY6 PRELIMINARY; PRT; 787 AA.
AC Q9VEY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SB gene product.
GN SB OR CG4316.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

```


Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.J., Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Mazzarelli J., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -I- SIMILARITY: CONTAINS 1 CUB DOMAIN
CC -I- SIMILARITY: BELONGS TO PETIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
CC EMBL; AK004939; BAB23684.1; -.
DR HSP; P00763; LDFO.
DR MEROPS; S01.308; -.
DR MGD; MGI:1919003; 1300008A22Rik.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00057; ldl_recept_a; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; P00722; CHYMOTRYPSIN.
DR PRINTS; P00261; LDLRECEPTOR.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00020; Tryp-Spc; 1.
DR PROSITE; PS00180; CUB; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS50068; LDLRA_2; 3.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Glycoprotein; Hydrolase; Serine protease.
KW SEQUENCE 799 AA; 89557 MW; 16315A646AD5288 CRC64;

Alignment Scores:		
Pred. No.:	7.28e-41	799
Score:	480.50	Matches: 96
Percent Similarity:	55.95%	Conservative: 45
Best Local Similarity:	38.10%	Mismatches: 88
Query Match:	28.81%	Indels: 23
DB:	11	Gaps: 5

US-09-735-713A-1 (1-921) x Q9DBI0 (1-799)

160	ACTCGCATCTTTGGAGCAACCCAAAGTGGAGAGGGTTCTATCCCTGGCAGGTATCTCTG	219
QY		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		
201		
202		
203		
204		
205		
206		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		
239		
240		
241		
242		
243		
244		
245		
246		
247		
248		
249		
250		
251		
252		
253		
254		
255		
256		
257		
258		
259		
260		
261		
262		
263		
264		
265		
266		
267		
268		
269		
270		
271		
272		
273		
274		
275		
276		
277		
278		
279		
280		
281		
282		
283		
284		
285		
286		
287		
288		
289		
290		
291		
292		
293		
294		
295		
296		
297		
298		
299		
300		
301		
302		
303		
304		
305		
306		
307		
308		
309		
310		
311		
312		
313		
314		
315		
316		
317		
318		
319		
320		
321</		

Db	662	LeuAspHisProValValTyr-SerAlaThrValArgProValCysLeuProAlaArgSer	681
QY	514	GAGCAATTTGAGCGTGTTTTATTGTGACAACTGCAGCGTGGGGCCCTTAACTGAGGT	573
Db	682	HisPheGluProGlyGlnHisCysTrpIleThrGlyTrpGlyAlaGlnArgGluGly	701
QY	574	GGCGTCCTCTCACAAAGTCTTCAGCAAGATGAATCTGCCTATTTTGACCTGGGAAGAGTG	633
Db	702	GlyProValSerAsnThrLeuGlnLysValAspValGlnLeuValProGlnAspLeuCys	721
QY	634	GTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTTCACAGGT	693
Db	722	SerGluAlaTyrArgTyrGlnValSerPro-----ArgMetLeuCysAlaGly	737
QY	694	TTTCTGTATGGAGGAGGAGACGATGTCCAGGGAGATTCAGGAGGTTTCACATGTGCCGG	753
Db	738	TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg	757
QY	754	AATAAGAAAGGGCGCTGGACTCTGGCTGGTGTGACTTCCCTGGGGTTGGGCTGTGTCGA	813
Db	758	GluProSerGlyArgTTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg	777
QY	814	GGCTGGGAAACAATGTGAGGAAAAGTGATCAAGGATCCCT-----GGGATCTTC	864
Db	778	-----ProAsnPheGlyValTyr	784
QY	865	ACACACATTAGTAAGTCTTTCTTGGATCCAGCA	900
Db	785	ThrArgValThrArgValIleAsnTrpIleGlnGln	796

RESULT 7
Q8VIF2

ID	Q8VIF2	PRELIMINARY;	PRT;	335 AA.
AC	Q8VIF2;			
DT	01-MAR-2002	(TrEMBLrel. 20, Created)		
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)		
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)		
DE	Testis serine protease.			

GN
TESSP2.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;
PN f11

KN [T]
RP SEQUENCE FROM N.A.

RA Matsui H., Takano N., Takahashi T.;

RT "Tessp2,serine protease specifically expressed in mouse testis.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBPJ databases

DR EMBL; AB052292; BAB78735.1; -;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00099. truncation.
DR InterPro: IPR001254; Ser_protease_Try.

DR PIAM; PF00089; trypsin; I.
DR PRINTS: PR00722: CHYMOTRYPSIN.

DR SMART; SM00020; Tryp_SPC; 1.

```
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
```

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN

DR FROSTIE, F300133; IRIFSIK_SER; UNKNOWN;
KW Protease.

SQ SEQUENCE 335 AA; 36682 MW; E8FC667

Alignment Scores:

alignment scores:
Pred. No.: 1.93e-40 Length

Score: 475.50

Percent Similarity:	54.91%	Cons
Rest Local Similarity:	37.00%	Misc

Best Local Similarity: 37.09%
Query Match: 28.51%

DB: 11 Gaps

115-09-735-713A-1 (1-021) v 08VFE2 (1-225)

US-09-733-1 (1-921) X Q8VIF2 (1-335)

QY 97 AAAGCTCCCAGTTGTGGCCAGAGCTCGGTTAAGGTACAGCCTTGGAAATTATTTAACATT 156

62 ARSENOPHOSPHOMOTACENOPHOSPHORIC ACID-DE

DD 62 ArgSerProPheMetAsnPheSerLeuValCysGlyGlnPro----- 75

QY 157 TTCAAGTCCATCTTGGAGAGCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCT 216
 DB 76 PheMetLysIleMetGlyGlyValAspAlaGluGlyLysTrpProTrpGlnValSer 95
 QY 217 CTGAACAAGAGGAGAGCATATTTGTGGAGGAGCATCTCTCACACAGTGGGTATC 276
 DB 96 ValArgValArgHisMetHisValCysGlyGlySerLeuIleAsnSerGlnTrpValLeu 115
 QY 277 ACCGGCGCTCAGTCATCTGCAACACAGAACATTTGTGTCTACTTTGAATGTTACTGCTGGA 336
 DB 116 ThrAlaAlaHisCysIleTyrSerArg-----IleGlnTyrAsnValLysValGly 132
 QY 337 GAGTATGACTTAAGCCAGCAGACAGCCAGGAGCAAACTCTCACTATTGAACCTGTGCATC 396
 DB 133 AspArgSerValTyrArgGlnAsnThr---SerLeuValIleProIleLysThrIlePhe 151
 QY 397 ATACATCCACATTTCTCCACCAAGAACCAATGAGCATATGATATGTCCTTTGAAGATG 456
 DB 152 ValHisProLysPheSerThrThrIleValValLysAsnAspIleAlaLeuLeuLysLeu 171
 QY 457 GCTGGAGCCTTCCAAATTTGCCACTTTGTGGGCCCATATGCTCTCCAGAGCTCGGGGAG 516
 DB 172 GlnHisProValAsnPheThrThrAsnIleTyrProValCysIleProSerGluSerPhe 191
 QY 517 CAATTTGAGGCTGGTTTATTTGTACAACTGCAGGCTGGGGCCGCTTAACTGAAGGTGGC 576
 DB 192 ProValLysAlaGlyThrLysCysTrpValThrGlyTrpGlyLysLeuValProGlyAla 211
 QY 577 -----GTCCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCTATTTGACCTGGGAAGAG 630
 DB 212 ProAspValProThrGluIleLeuGlnGluValAspGlnAsnValIleLeuTyrGluGlu 231
 QY 631 TGTGTGCAGCTCTGTTAACTAAAGAGGCCATCAGTGGG-----AAG 675
 DB 232 CysAsnGlu-----MetLeuLysLysAlaThrSerSerValAspLeuValLys 248
 QY 676 ACCTTTCTTCCACAGGTTTCTCTGATGGAGGAGAGCAGCATGTCTCAGGAGATTCAGGA 735
 DB 249 ArgGlyMetValCysGlyTyrLysGluArgGlyLysAspAlaCysGlnGlyAspSerGly 268
 QY 736 GGTCTACTCATGTC-----CGGATAAGAAAGGGCCCTGGAGCTGGCTGGTGTGCTACT 789
 DB 269 GlyProMetSerCysGluPheGluAsnLys-----TrpValGlnValGlyValVal 285
 QY 790 TCTCGGGGTTGGGCTGTGTGTGGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGA 849
 DB 286 SerTrpGlyIleSerCysGlyArg-----LysGly 295
 QY 850 TCCCTGGGATCTTCAACAGACATTAAGTAAAGTGTCTTCTCTGGATC 894
 DB 296 TyrProGlyValTyrThrAspValAlaPheTyrSerLysTrpLeu 310

RESULT 8

Q9DAT3 ID Q9DAT3 PRELIMINARY: PRT; 624 AA.
 AC Q9DAT3;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DE 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE 1600027G01Rik protein (RIKEN cDNA 1600027G01 gene).
 GN 1600027G01Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Sakurai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gystincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AK005546; BAB24114.1; -;
 DR EMBL; BC019485; AAH19485.1; -;
 DR HSSP; P00750; 1RTF.
 DR MEROPS; S01.213; -;
 DR MGD; MGI:1919281; 1600027G01Rik.
 DR InterPro; IPR000177; Apple.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR001254; Ser--protease_Try.
 DR Pfam; PF00024; PAN; 4.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00005; APPLIEDOMAIN.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00223; APPLP; 4.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00495; APPLE; UNKNOWN_3.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 624 AA; 69788 MW; 0BEDDEBC56009E97 CRC64;
 Alignment Scores:
 Pred. No.: 3.65e-40 Length: 624
 Score: 473.50 Matches: 95
 Percent Similarity: 53.82% Conservative: 46
 Best Local Similarity: 36.26% Mismatches: 98
 Query Match: 28.39% Indels: 23
 DB: 11 Gaps: 6
 US-09-735-713a-1 (1-921) x O9DAT3 (1-624)
 QY 127 AAGTACAGCCTTGGAAATATTTTAAACATTTTCACTGCGCATTTTGGAGGAGCAAGTG 186
 DB 385 LysIleAsnPro-----ArgValValGlyGlyAlaAlaSer 396
 QY 187 GAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGAGCAGCATATTGTGGGA 246
 DB 397 ValHisGlyGluTrpProTrpGlnValThrLeuHisIleSerGlnGlyLeuLysCysGly 416
 QY 247 GGAAGCATCTCTCACCACAGTGGGTATCATCAGCGGCTCTACTGTCATTCGAAACAGAAC 306
 DB 417 GlySerIleIleGlyAsnGlnTrpIleLeuThrAlaAlaHisCysPheSerGlyIleGlu 436
 QY 307 ATTGTGCTACTTTGAATGTACTGCTGGAGAGTATGACTTAAAGCAGACAGACCCAGGA 366
 DB 437 ThrProLysLysLeuArgValTyrGlyGlyIleValAsnGlnSerGluIleAsnGly 456
 QY 367 GAGCAAACTCTCACTATTGAACCTGTCTATCATATCATATTTCTCCCAAGAAACCA 426

```

Db 457 ThrAlaPhePheArgValGlnGluMetIleIleHisAspGlnTyrThrAlaGlu--- 475
QY 427 ATGGACTATGATATGCGCTTTTGAAGATGCTGGAGCGCTTCCAATTTGGCCACTTTGTG 486
Db 476 SerGlyTyrAspIleAlaLeuLeuLysLeuGluSerAlaMetAsnTyrThrAspPheGln 495
QY 487 GGGCCCATATCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTATTTGACAACT 546
Db 496 ArgProIleCysLeuProSerLysGlyAspArgAsnAlaValHisThrGluCysTrpVal 515
QY 547 GCAGGCTGGGCCCTTAACTGAAGGTGGCTCTCTCACAAGTCTTGCAGGAAGTGAAT 606
Db 516 ThrGlyTrpGlyTyrThrAlaLeuArgGlyGluValGlnSerThrLeuGlnLysAlaLys 535
QY 607 CTGCTATTTTGACCTGGGAGAGCTGTGGCAGCTCTGTTAACACTAAAGAGGCCCATC 666
Db 536 ValProLeuValSerAsnGluGluCysGlnThrArgTyr-----ArgAlaHisLysIle 553
QY 667 AGTGGGAAGACCTTTCTTTCACAGGTTTTCCTGATGAGGGAGAGAGCGCATGTCAGGA 726
Db 554 ThrAsnLys---MetIleCysAlaGlyTyrLysGluGlyLysAspThrCysLysGly 572
QY 727 GATTCAGAGGTTCTACTGTCGCGGAATAAGAAAGGGCGCTGCTGCTGGTGGTGTG 786
Db 573 AspSerGlyGlyProLeuSerCysLys---TyrAsnGlyValTrpHisLeuValGlyIle 591
QY 787 ACTTCCTGGGTTTGGCTGTGTCGAGGCTGGAGGCTGGAGAAACAATGTGAGAAAGTATCAA 846
Db 592 ThrSerTrpGlyGluGlyCysGlyGlnLysGluArg----- 603
QY 847 GGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTTCCCTGGATCCACGAGCAATC 906
Db 604 -----ProGlyValTyrThrAsnValAlaLysTyrValAspTrpIleLeuGluLysThr 621
QY 907 CAAACT 912
Db 622 GlnThr 623

RESULT 9
Q91V47 PRELIMINARY; PRT; 624 AA.
AC Q91V47;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor XI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.WR; TISSUE=LIVER;
RX MEDLINE=97385041; PubMed=9242536;
RT Gallani D., Sun M.F., Sun Y.;
RT "A comparison of murine and human factor XI.";
RL Blood 90:1055-1064(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.WR; TISSUE=LIVER;
RA Gallani D., Sun M.F., Sun Y.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF356627; AAK40233.1;
DR MEROPS; S01.213;
DR InterPro; IPR000177; Apple.
DR InterPro; IPR003014; PAN.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00495; APPLE; UNKNOWN.2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN.1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN.1.
SQ SEQUENCE 624 AA; 69874 MW; 49D281BFAEC12A03 CRC64;
Alignment Scores:
Pred. No.: 1.53e-39 Length: 624
Score: 467.50 Matches: 95
Percent Similarity: 53.44% Conservative: 45
Best Local Similarity: 36.26% Mismatches: 99
Query Match: 28.03% Indels: 23
DB: 11 Gaps: 6
US-09-735-713a-1 (1-921) x Q91V47 (1-624)
QY 127 AAGGTACAGCCTTGGGAATTATTTTAAACATTTTCAGTCGCATTCTTGAGAGAACCAAGTGT 186
Db 385 LysIleAsnPro-----ArgValValGlyGlyAlaAlaSer 396
QY 187 GAGAAGGTTCCCTATCCCTGGCAGGTATCTCTGAAACAAAGCGAGAGCATATTTGTGGA 246
Db 397 ValHisGlyGluTrpProTrpGlnValThrLeuHisIleSerGlnGlyHisLeuCysGly 416
QY 247 GGAAGCATCTCTCACCACAGTCGGGTGATCACCAGCGCTCCTGCTGCAATTCGCAAAACAGAAAC 306
Db 417 GlySerIleIleGlyAsnGlnTrpIleLeuThrAlaAlaHisCysPheSerGlyIleGlu 436
QY 307 ATTGTCTACTTGTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGACCCAGGA 366
Db 437 ThrProLysLysLeuArgValTyrGlyIleValAsnGlnSerGluIleAsnGluGly 456
QY 367 GAGCAACTCTCACTATTGAACTGCATCATCATATCATCTTCCACCAAGAAACCA 426
Db 457 ThrAlaPhePheArgGluGlnGluMetIleIleHisAspGlnTyrThrAlaGlu--- 475
QY 427 ATGGACTATGATATGCGCTTTTGAAGATGCTGGAGCGCTTCCAATTTGGCCACTTTGTG 486
Db 476 SerGlyTyrAspIleAlaLeuLeuLysLeuGluSerAlaMetAsnTyrThrAspPheGln 495
QY 487 GGGCCCATATGTTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTGTATTTTGTACAACT 546
Db 496 ArgProIleCysLeuProSerLysGlyAspArgAsnAlaValHisThrGluCysTrpVal 515
QY 547 GCAGGCTGGGGCGCTTAACTGAAGTGGCGCTCTCACAAGTCTTGCGAGAAAGTGAAT 606
Db 516 ThrGlyTrpGlyTyrThrAlaLeuArgGlyGluValGlnSerThrLeuGlnLysAlaLys 535
QY 607 CTGCTATTTTGAACCTGGGAAGAGTGTGGCAGCTCTGTTAACACTAAAGAGGCCCATC 666
Db 536 ValProLeuValSerAsnGluGluCysGlnThrArgTyr-----ArgAlaHisLysIle 553
QY 667 AGTGGGAAGACCTTTCTTTCACAGGTTTTCCTGATGAGGGAGAGACGCATGTGAGGA 726
Db 554 ThrAsnLys---MetIleCysAlaGlyTyrLysGluGlyGlyLysAspThrCysLysGly 572
QY 727 GATTCAGAGGTTCTACTGTCGCGGAATAAGAAAGGGCGCTGCTGCTGGTGGTGTG 786
Db 573 AspSerGlyGlyProLeuSerCysLys---TyrAsnGlyValTrpHisLeuValGlyIle 591
QY 787 ACTTCCTGGGTTTGGGCTGTGTCGAGGCTGGAGCAACAATGTGAGAAAGTATCAA 846
Db 592 ThrSerTrpGlyGluGlyCysGlyGlnLysGluArg----- 603
QY 847 GGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTTCCCTGGATCCACGAGCAATC 906
Db 604 -----ProGlyValTyrThrAsnValAlaLysTyrValAspTrpIleLeuGluLysThr 621
QY 907 CAAACT 912
Db 622 GlnThr 623
RESULT 10
Q967X8 PRELIMINARY; PRT; 467 AA.
ID Q967X8

```



```

AC Q967X8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CUB-serine protease.
OS Panulirus argus (Spiny lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;
OC Palinuroidea; Palinuridae; Panulirus.
OX NCBI_TaxID=6737;
RN [1]
RP SEQUENCE FROM N.A.
RA Levine M.Z., Maithall W.W., Tai P.C., Derby C.D.;
RT "Molecular cloning, characterization, cellular localization and
RT possible function of a CUB-serine protease in the olfactory system of
RT the spiny lobster Panulirus argus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF357226; AAK48984.1; -
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00431; CUB; 1.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE: PS01180; CUB; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS0135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 467 AA; 50453 MW; 1D2E663D3314BBFD CRC64;

Alignment Scores:
Pred. No.: 4.7e-39 Length: 467
Score: 462.50 Matches: 105
Percent Similarity: 52.73% Conservative: 40
Best Local Similarity: 38.18% Mismatches: 97
Query Match: 27.73% Indels: 33
DB: 5 Gaps: 9

US-09-735-713a-1 (1-921) x Q967X8 (1-467)
QY 103 CCAGTTGTGGCAGAGCTGGTTAAGGTACAGCTTGGAAATTTTAACTTTTCAGT 162
Db 216 ProSerCysAlaCysGlyAsnVal-----AsnArgAlaThr 227
QY 163 CGCATCTTGGAGGAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCGAAA 222
Db 228 ArgIleValGlyGlnGluThrGluValAsnGluTyrProTrpGlnValLeuVal 247
QY 223 CAAGGCGAAGCAT---ATTGTGGAGGAGCATGCTCTCACACAGTGGGTATCAGC 279
Db 248 ThrArgAspMetTyrValIleCysGlySerIleSerSerGlnTrpValLeuThr 267
QY 280 GCGGCTCACTGCATTCGCAACAGAAATTTGTCTACTTTGAATGTTACTGCTGGAGAG 339
Db 268 AlaAlaHisCysValAspGlyGlyAsnIleGlyTyrValLeu-----ValGlyAsp 284
QY 340 TATGACTTAAGCCAGACAGACCAGGAGAGCAAACTCTACTATTGAACTGTC----- 393
Db 285 HisAsnPheAlaSerThrAspAspThrThrThrSerArgLeuValGluValGlnIle 304
QY 394 ATCATCATCCACATTTCTCCACAGAAACCAATGAGCATATGATATGTCCTTTTGAAG 453
Db 305 IleSerHisProAspTyrAspSer---SerThrValAspAsnAspMetAlaLeuLeuArg 323
QY 454 ATGGCTGGAGCCCTCCCAATTTGGCCACTTTGTGGGGCCCATATGTTCTCCAGAGTCGG 513
Db 324 LeuGlyGluAlaLeuGluPheThrArgGluValAlaProValCysLeuProSerAsnPro 343
QY 514 GAGCAATTTGAGGCTGTTTATTATTGTACAACTGCAGCTGGGCGCCGCTTAACTGAAGGT 573
Db 344 ThrGluAspTyrAlaGlyValThrAlaThrValThrGlyTyrGlyAlaThrThrGluGly 363
QY 574 GCGTCTCTCACAAAGTCTTTCAGGAGTGAATCTGCTATTTTGGACCTGGGAAGAGTGT 633
Db 170 ValSerIleAsnHisLeuLeuSerAspLysValThrAlaLeuHisHisSerValTyr 189

```

```

* Db 364 GlySerMetSerValThrLeuGlnGluValAspValProValLeuThrThrAlaAlaCys 383
QY 634 GTGGCAGCTCTGTAAACACTAAAGAGGCCCATCAGTGGGAGACCTTTCTTTGCACAGGT 693
Db 384 SerSerTrpTyrSerSerLeuThr-----AlaAsnMetMetCysAlaGly 398
QY 694 TTTCTCTGATGGAGGAGACGATGTCAGGAGATTCAGGAGGTTTCACATGTCGCCGG 753
Db 399 PheSerAsnGluGlyLysAspSerCysGlnGlyAspSerGlyGlyPro---MetValTyr 417
QY 754 AATAAGAAAGGGCGCTGGACTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 813
Db 418 SerAlaThrSerAsnTyrGluGlnIleGlyValValSerTrpGlyArgGlyCysAlaArg 437
QY 814 GGCTGGAGAAACAATGTGAGGAAAGTATCAAGAGATCCCTGGATCTTCACAGACATT 873
Db 438 -----ProGlyPheProGlyValTyrAlaArgVal 447
QY 874 AGTAAAGTGTCTTCTGATGATCCAGCAACATCCAACTGGTAAAC 918
Db 448 ThrGluTyrLeuGluTrpIleAlaAlaAsn-----ThrGlyAsn 460

RESULT 11
Q8VDE0 PRELIMINARY; PRT; 453 AA.
AC Q8VDE0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE TMPSR33 protein.
GN TMPSR33.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Guipponi M., Scamuffa N., Scott H.S., Rossier C., Antonarakis S.E.;
RT "Isolation and characterization of the mouse Tmprs3 gene.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ300738; CAC83350.1; -.
DR MGD; MGI:2155445; Tmprs3.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001190; Srcr_receptor.
DR Pfam: PF00057; ldl_recept_a; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; UNKNOWN_1.
DR PROSITE; PS50068; LDLRA_2; 1.
DR PROSITE; PS50287; SRCR_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ SEQUENCE 453 AA; 49529 MW; 21E5697DC8781BD3 CRC64;

Alignment Scores:
Pred. No.: 5.25e-39 Length: 453
Score: 462.00 Matches: 102
Percent Similarity: 51.62% Conservative: 57
Best Local Similarity: 33.12% Mismatches: 115
Query Match: 27.70% Indels: 34
DB: 11 Gaps: 8

US-09-735-713a-1 (1-921) x Q8VDE0 (1-453)
QY 1 ATGAGTCTCAAAATGCTTATTAAGCAGCAACAGCTGATTTTACTAGGAATAGTCTTT 60
Db 170 ValSerIleAsnHisLeuLeuSerAspLysValThrAlaLeuHisHisSerValTyr 189

```

```

QY 61 TTTGACRAGGT---AAATCTGCARCTCTTTTCGTCCTCCCAAGCTCCAGCTGTGGCCAG 117
Db 190 MetArgGluGlyCysThrSerGlyHisValValThrLeuLysCysSerAlaCysGlyThr 209
QY 118 AGTCGTGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTCAGTCGCATTTCTTGGAGGA 177
Db 210 ArgThr-----GlyTyrSerProArgIleValGlyGly 220
QY 178 AGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGGCAGACAT 237
Db 221 AsnMetSerSerLeuThrGlnTrpProTrpGlnValSerLeuGlnPheGlnGlyTyrHis 240
QY 238 ATTTGTGGAGAAAGCATCGCTCACACAGCTGGGTATCACGCGGCTCATCTCATGTGA 297
Db 241 LeuCysGlyGlySerIleIleThrProLeuTrpIleValThrAlaAlaHisCysValTyr 260
QY 298 AACAGAAACATTTGCTACTTGAATTTACTGCTGGAGAGTATGACTTTAAGCCAGACA 357
Db 261 AspLeuTyrHisProLysSerTrpThrValGlnValGlyLeu-----ValSerLeuMet 278
QY 358 GACCCAGGAGCAAACTCCTACTATTGAACTGTCTATCATATCATCATCTTCTCCACC 417
Db 279 AspSerProValProSerHisLeuValGluLysIleIleTyrHisSerLysTyrLysPro 298
QY 418 AAGAAACCAATGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCTTCCCAATTTGGC 477
Db 299 LysArg---LeuGlyAsnAspIleAlaLeuMetLysLeuSerGluProLeuThrPheAsp 317
QY 478 CACTTTGTGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCTGGTTTTATT 537
Db 318 GluThrIleGlnProIleCysLeuProAsnSerGluLysAsnPheProAspGlyLysLeu 337
QY 538 TGTACAACTCGAGCTGGGGCCCTTAACCTGAAGTGGCGTCTCTCAACAAGCTTTCGAG 597
Db 338 CysTrpThrSerGlyTrpGlyAlaThrGluAspGlyGlyAspAlaSerProValLeuAsn 357
QY 598 GAAGTGAATCTGCCTATTTTGAACCTGGGAGAGTGTGGCAGCTCTGTAAACACTAAAG 657
Db 358 HisAlaAlaValProLeuIleSerAsnLysIleCys-----AsnHis 371
QY 658 AGCCCATCAGTGGG-----AAGACCTTTCTTTGACAGGTTTCTCTGATGGA 705
Db 372 ArgAspValTyrGlyGlyIleIleSerProSerMetLeuCysAlaGlyTyrLeuLysGly 391
QY 706 GGGAGAGCGCATGTCCAGGAGATTCAGGAGTTCACATCATCTGCGGGAATAAGAAAGG 765
Db 392 GlyValAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGluArg--- 410
QY 766 GCCTGGACTCTGGCTGGTGTGACTTCTCTGGGGTTTGGGCTGTGGTGTGGAGGCTGGAGAAC 825
Db 411 LeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys----- 424
QY 826 AATGTGAGGAAAGTATCAAGGATCCCTGGATCTTCACAGACATTAAGTAAAGTGCTT 885
Db 425 -----AlaGluValAsnLysProGlyValTyrThrArgIleThrSerPheLeu 440
QY 886 TCCTGGATCCAGCAACATCAAA 909
Db 441 AspTrpIleHisGluGlnLeuGlu 448
RESULT 12
Q8R0P5 PRELIMINARY; PRT; 638 AA.
AC Q8R0P5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_faxid=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026555; AAH26555.1; -.
SQ SEQUENCE 638 AA; 71382 MW; CC27C93F4B57C599 CRC64;

Alignment Scores:
Pred. No.: 1.49e-38 Length: 638
Score: 458.00 Matches: 94
Percent Similarity: 53.10% Conservative: 43
Best Local Similarity: 36.43% Mismatches: 103
Query Match: 27.46% Indels: 18
DB: 11 Gaps: 6

US-09-735-713a-1 (1-921) x Q8R0P5 (1-638)
QY 154 ATTTTCAGTCGCATTTCTTGAGGAAGCCCAAGTGGAGAGGTTTCCTATCCCTGGCAGGTA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGlyGluTrpProTrpGlnVal 406
QY 214 TCTCTCAAAACAAGG-----CAGAAGCATATTGTGGAGGAAGCATCTCTCAACCA 264
Db 407 SerLeuGlnValLysLeuValSerGlnThrHisLeuCysGlyGlySerIleIleGlyArg 426
QY 265 CAGTGGGTGATCACGGCGCTCAGTCATTCGCAATTCGCAACAGAAACATTTGCTACTTTGAAT 324
Db 427 GlnTrpValLeuThrAlaAlaHisCysPheAspGlyIleProTyrProAspValTrpArg 446
QY 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGACAGCCAGAGAGCAAACTCTCACTATT 384
Db 447 IleTyrGlyGlyIleLeuSerLeuSerGluIleThrLysGluThrProSerSerArgIle 466
QY 385 GAAACTGTCATCATACATCCACATTTCTCCACCAGAAACCAATGACTATGATATTGCC 444
Db 467 LysGluLeuIleIleHisGlnGluTyrLysValSerGluGly---AsnTyrAspIleAla 485
QY 445 CTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTTCCA 504
Db 486 LeuIleLysLeuGlnThrProLeuAsnTyrThrGluPheGlnLysProIleCysLeuPro 505
QY 505 GAGCTCGGGGACATTTGAGGCTGGTATTATTGTACAACTGCAGGCTGGGCGCGCTTA 564
Db 506 SerLysAlaAspThrAsnThrIleTyrThrAsnCysTrpValThrGlyTrpGlyTyrThr 525
QY 565 ACTGAAGTGGCGCTCTCTCACAAGTCTTGCAAGAGTGAATCGCTCTATTTTGGACCTGG 624
Db 526 LysGluGlnGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
QY 625 GAAGAGTGTGGCGAGCTCTGTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTT 684
Db 546 GluGluCysGlnLys-----LysTyrArgAspTyrValIleAsnLysGlnMetIle 562
QY 685 TGCACAGCTTTCTCTGATGGAGGAGACACGATCTCAGGAGATTCAGGAGTTTCATCTC 744
Db 563 CysAlaGlyTyrLysGluGlyGlyThrAspAlaCysLysGlyAspSerGlyGlyProLeu 582
QY 745 ATGTGCCGGAATAACAAAGGCGCTCTGGACTCTGGCTGTGTGACTTCTCTGGGTTGGGC 804
Db 583 ValCysLysHis---SerGlyArgTrpGlnLeuValGlyIleThrSerTrpGlyGluGly 601
QY 805 TGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAAGTGTGAAGATCCCTCGGATCTTC 864
Db 602 CysAlaArg-----LysaspGln-----ProGlyValTyr 611
QY 865 ACACACATTAGTAAGTCCTTCTCGATCCAGCAACACATCCCAACACTGGTAAAC 918
Db 612 ThrLysValSerGluTyrMetAspTrpIleLeuGluLysThrGlnSerSerAsp 629

RESULT 13
Q99L44
ID Q99L44 PRELIMINARY; PRT; 339 AA.

```

AC Q99L44;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to protease, serine, 8 (Prostasin).
 GN PRSS8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ, AND SWISS; TISSUE=LUNG;
 RA Vergnese G.W., Caughey G.H.;
 RT "Molecular cloning and characterization of mouse prostasin, a type I
 RT membrane-associated serine protease of the gamma-tryptase/prostasin
 RT gene family.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Kitamura K., Takefumi N., Kimio T.;
 RT "mouse serine protease.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; BC003851; AAH03851.1; -;
 DR EMBL; AF378086; AAL06320.1; -;
 DR EMBL; AF378085; AAL06319.1; -;
 DR EMBL; AB038244; BAB82496.1; -;
 DR HSSP; P00734; IUVS.
 DR MEROPS; S01.159; -;
 DR MGD; MGI:1923810; Prss8.
 DR InterPro; IPR001314; Chymotrypsin.
 DR Pfam; PF00089; trypsin_1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP-SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 339 AA; 36216 MW; BC2DE88BC057AF10 CRC64;

Alignment Scores:
 Pred. No.: 8.59e-38 Length: 339
 Score: 450.00 Matches: 100
 Percent Similarity: 48.15% Conservative: 43
 Best Local Similarity: 33.67% Mismatches: 120
 Query Match: 26.98% Indels: 34
 DB: 11 Gaps: 8

US-09-735-713A-1 (1-921) x Q99L44 (1-339)

QY 34 CTGATTTTACTAGTAGTCTTTTGTGAACRAGGTAAATCTGCARCTCTTTGCTC 93
 Db 17 IleLeuLeuLeuLeuGlyLeuLeu-----GlnSerGlyIleArgAlaAsp 31
 QY 94 CCAAAAGTCCCAAGTTGTGTGGCAGAGCTGCTGTTAAGGTACAGCTTGGAAATATTTTAAAC 153
 Db 32 GlyThrGluAlaSerCysGly-----AlaValIleGlnPro----- 43
 QY 154 ATTTTCAGTCGCATTTCTGGAGGAGCCCAAGTCGAGAGGGTTTCCTCCCTGGCAGGTA 213
 Db 44 -----ArgIleThrGlyGlyGlySerAlaLysProGlyGlnTrpProTrpGlnVal 60
 QY 214 TCTCTGAACAAAGGAGAGCATATTTTGGAGGAGCATCTCTCCACCAAGTGGGTG 273
 Db 61 SerIleThrTyrAspGlyAsnHisValCysGlyGlySerLeuValSerAsnLysTrpVal 80

QY 274 ATCAGCGCGCTCAGTCGATTCGAAACAGAAACATCTGTCTACTTTGAATGTACTGCT 333
 Db 81 ValSerAlaAlaHisCysPheProArgGluHisSerArgGluAlaTyrGluValLysLeu 100
 QY 334 GGAGACTATGACTTAAGCCAGACAGACCAGGAGAGCAAACTCTCACTATTGAACGTGC 393
 Db 101 GlyAlaHisGlnLeuAspSerTyrSerAsnAspThrValValHisThrValAlaGlnIle 120
 QY 394 ATCATATCATCCCACTTCTCCACCAAGAACCAATGCATATATATATGCTTCCAGAGCTCGG 453
 Db 121 IleThrHisSerSerTyr---ArgGluGluGlySerGlnGlyAspIleAlaLeuIleArg 139
 QY 454 ATGCTCTGGAGCCCTCCAAATTTTGTGGGCCCATATCTCTTCCAGAGCTCGCG 513
 Db 140 LeuSerSerProValThrPheSerArgTyrIleArgProIleCysLeuProAlaAlaAsn 159
 QY 514 GACCAATTTGAGGCTGGTTTATTTGTACAACTGCAGGCTGGCGGCGCTTAACTGAAGGT 573
 Db 160 AlaSerPheProAsnGlyLeuHisCysThrValThrGlyTrpGlyHisValAlaProSer 179
 QY 574 GCGCTCTCTCA-----CAAGTCTTGCAGGAAGTGAATCTGCTATTTTGACCTGGAA 627
 Db 180 ValSerLeuGlnThrProArgProLeuGlnGlnLeuValProLeuIleSerArgGlu 199
 QY 628 GAGTGTGTGCGAGCTCTGTTAACACTAAAGAGCGCCATCAGTGGGAAGACC----- 678
 Db 200 ThrCysSerCysLeuTyrAsnIleAsnAlaValProGluGluProHisThrIleGlnGln 219
 QY 679 ---TTTCTTTCACAGGTTTCTGATGGAGGAGGAGCAGCATGTCTCAGGAGATTCAGGA 735
 Db 220 AspMetLeuCysAlaGlyTyrValLysGlyLysAspAlaCysGlnGlyAspSerGly 239
 QY 736 GGTTCACATGTGCGCGAATAAGAAGGCGCTGCACTCTGCTGGTGTGCTGCTCTGG 795
 Db 240 GlyProLeuSerCys---PrometGluGlyIleTrpTyrLeuAlaGlyIleValSerTrp 258
 QY 796 GGTTTGGGCTGTGCTGAGGCTGGAGAACAAATGTGAGGAAAAGTATCAAGGATCCCT 855
 Db 259 GlyAspAlaCysGlyAlaProAsnArg-----Pro 268
 QY 856 GGGATCTTCACAGACATTAGTAAAGTCTTCTGTCGATCCACGACACATC 906
 Db 269 GlyValTyrThrLeuThrSerThrTyrAlaSerTrpIleHisHisVal 285

RESULT 14
 O97506 PRELIMINARY; PRT; 643 AA.
 ID O97506;
 AC O97506;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
 DE Kallikrein.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=PLASMA;
 RA Kimura A., Kimura T., Okimura H., Hamabata T., Ohnishi J.,
 RA Moriyama A., Takahashi K., Takahashi T.;
 RT "Identification of porcine follipain as plasma kallikrein, and its
 RT possible involvement in the production of bradykinin within the
 RT follicles of porcine ovaries.";
 RL Mol. Reprod. Dev. 57:79-87(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AB022425; BAA37147.1; -;
 DR HSSP; P00766; ICHG.
 DR MEROPS; S01.212; -;
 DR InterPro; IPR000177; Apple.
 DR InterPro; IPR001314; Chymotrypsin.


```

Db 241 LeuLys---ValLysPheLeuArgIleValAlaGlyGluHisAspLeuGluValAspGlu 259
QY 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCACTACATCCACATTTCTCCACCAAG 420
Db 260 GlyThrGluGlnLeuIleGlnValAspGlnMetPheThrHisProAlaTyr---ValSer 278
QY 421 AAACCAATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGCCAC 480
Db 279 GluThrAlaAspSerAspIleAlaLeuLeuArgLeuArgThrProIleValTyrSerVal 298
QY 481 TTTGTGGGGCCCATATCTTCCAGAGCTGCGGGAGCAATTTGAGCTGGTCTGTTTATTGT 540
Db 299 TyrAlaValProValCysLeuPro---LeuArgGluMetAlaGluArgGluLeuTrpAla 317
QY 541 -----ACAACTGCAGCTGGCGCTTAACTGAAGTGGCGCTCTCTCCACAA 588
Db 318 ValSerLysHisThrValSerGlyTrpGlyLysArgSerGluAspGlyProThrSerArg 337
QY 589 GTCTTCGAGGAAGTGAATCTGCTATTTTGACCTGGGAAGAGTGTGTG---GCAGCTCTG 645
Db 338 LeuLeuArgArgLeuLeuValProArgIleArgThrGlnGluCysValGlnValSerAsn 357
QY 646 TTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTCTCTGATGGA 705
Db 358 LeuThrLeuThr-----SerAsnMetPheCysAlaGlyTyrIleGluGly 372
QY 706 GGGAGAGACCGATGTCAGGAGATTTCAGGAGTTTCACCTGTCGCGGAATAGAAAGG 765
Db 373 ArgGlnAspSerCysLysGlyAspSerGlyGlyProLeuValThrArgTyrArgAspThr 392
QY 766 GCGTGGACTCTGGCTGGTGTGACTTCCTGGGTTTGGGTTGGTGTGTCGAGGCTGGAGAAC 825
Db 393 AlaPhe---LeuLeuGlyIleValSerTrpGlyLysGlyCysAlaArg-----407
QY 826 AATGTGAGAAAGTGAATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTCTT 885
Db 408 -----ProGlySerTyrGlyIleTyrThrArgValSerAsnTyrLeu 421
QY 886 TCCTGGATCCAGCAACACATCCAAACT 912
Db 422 GlnTrpIleArgGlnThrThrAsnThr 430

```

Search completed: March 25, 2003, 07:04:11
Job time : 97.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 04:15:25 ; Search time 1854 Seconds
(without alignments)
8045.331 Million cell updates/sec

Title: US-09-735-713a-1

Perfect score: 921

Sequence: 1 atgagctctcaaaatgcttat.....acatccaaactggttaactaa 921

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479.8	52.1	670	10	BB625475
2	110.8	12.0	582	17	A2319576
3	104.4	11.3	585	17	A2320103
4	80.4	8.7	2246	11	AK005546
5	76.6	8.3	596	9	AL633968
6	76.4	8.3	675	13	BM104667

7	76.4	8.3	679	13	BM104492
8	74.8	8.1	627	13	BM104591
9	74.8	8.1	642	14	BM890103
10	74.8	8.1	642	14	BQ258513
11	74.8	8.1	645	13	BM183510
12	74.8	8.1	652	13	BM102995
13	74.8	8.1	653	13	BM157100
14	74.8	8.1	656	13	BM157053
15	74.8	8.1	657	14	BQ075190
16	74.8	8.1	658	14	BM860233
17	74.8	8.1	664	14	BM957767
18	74.8	8.1	668	14	BM861171
19	74.8	8.1	676	14	BM958270
20	74.8	8.1	682	14	BM777442
21	74.2	8.1	532	13	BM315903
22	74.2	8.1	540	14	BM860769
23	73.2	7.9	574	13	BM102953
24	73.2	7.9	681	14	BM957651
25	72.2	7.8	688	10	AW342995
26	72.2	7.8	689	13	BJ093907
27	71.6	7.8	592	13	BI984661
28	71.6	7.8	597	13	BI877311
29	71.6	7.8	673	14	BQ284373
30	71.2	7.8	732	10	AW154001
31	71.2	7.7	553	10	AW421919
32	70.6	7.7	638	10	AW153658
33	70.6	7.7	723	10	AW174475
34	70.4	7.6	557	9	AL595526
35	70	7.6	573	13	BM102765
36	70	7.6	708	10	AW153684
37	67	7.3	654	14	BM860845
38	66.2	7.2	432	13	BJ068921
39	64.4	7.0	564	10	AW153439
40	63.4	6.9	536	9	AI722029
41	62.2	6.8	538	10	AW342956
42	62.2	6.8	619	13	BI875854
43	62.2	6.8	3030	11	AK004939
44	61.8	6.7	504	9	AI721818
45	61.6	6.7	513	10	AW422658

ALIGNMENTS

RESULT 1 BB625475 670 bp mRNA linear EST 26-OCT-2001
LOCUS BB625475 RIKEN full-length enriched, adult male epididymis Mus
DEFINITION musculus cDNA clone 9230106D23 5', mRNA sequence.

ACCESSION BB625475.1 GI:16463737

VERSION BB625475

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 670)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,

Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,

M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki

D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,

Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

BM104492	fv48a09.y
BM104591	fv49b05.y
BM890103	fy64h10.y
BQ258513	fz82f02.y
BM183510	fv64d10.y
BM102995	fv39f04.y
BM157100	fv44c08.y
BM157053	fv43q06.y
BQ075190	fv43a08.y
BM860233	fy35q06.y
BM957767	fy69a06.y
BM861171	fy46g07.y
BM958270	fy71d06.y
BM777442	fy30h11.y
BM315903	fw66d06.y
BM860769	fy50d06.y
BM102953	fv39a08.y
BM957651	fy67f12.y
BJ093907	fw72a03.y
BI984661	fu05e10.y
BI877311	fl83f04.y
BQ284373	fz89d09.y
AW154001	fi29c04.y
AW421919	fi50f04.y
AW153658	fi24c05.y
AW174475	fi43f06.y
AL595526	AL595526
BM102765	fu31h10.y
AW153684	fi24e08.y
BM860845	fy51c04.y
BJ068921	BJ068921
AW153439	fi20f11.y
AI722029	fd18e05.y
AW342956	fi71b12.y
BI875854	fi63h07.y
AK004939	Mus muscu
AI721818	fd16g04.y
AW422658	fi45c12.y

Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES source

```

Location/Qualifiers
1..670
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9230106D23"
/clone_lib="RIKEN full-length enriched, adult male
epididymis"
/sex="male"
/tissue_type="epididymis"
/lab_host="adult"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTTTTATTTTATTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end:
BamHI"
BASE COUNT 180 a 154 c 169 g 167 t
ORIGIN
Query Match 52.18; Score 479.8; DB 10; Length 670;
Best Local Similarity 82.5%; Pred. No. 4.6e-137;
Matches 550; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 125 TTAAGGTACACCCCTTGGAAATTTATTAACATTTTCAGTCGCATCTTGGAGGAAGCCAAAG 184
||||| || ||| ||||||| ||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2 TTAAGCCCCACCTTCAGAAATTTACTTTAGCCCTTTTCAGTCGCATCTTGGAGGAAGCCAAAG 61
QY 185 TGGAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAGAGCAGAGCATATTTTGT 244
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 62 TGGAAAAGGGTTCCTATCTCTTGGCAAGTGTCTCTGAAACAAGAGCAGACATCTGTG 121
QY 245 GAGGAGGATCGTCTCACCACAGTGGGTGATCAGCGGGCTCAGTGCATGTCACAAACAGAA 304

```

```

122 GAGGAGCATTATCTCTTCACAGTGGGTGCATCACAGCGCTCATTTGCTAGCTTAACAGAA 181
QY 305 ACATTGTGCTTACTTGAATGTTACTGCTGGAGAGTAGTACTTAAGCCAGACACACCAG 364
182 ACATTGATTAACATTTTGAATGTAATCTGCTGGAGAACATGACTTGAGCCAGCCAGCCAG 241
QY 365 GAGAGCAAACTCTCATTATTTGAAACTGTTCATCATACATATCCACATTTCTCCACCAAGAAAC 424
Db 242 GAGAGCAAAACACTTGGCCATTGAAACCATCATCATACACCCCGCAGTTTCTCCACCAAGAAAC 301
QY 425 CAATGGACTATGATPATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTG 484
Db 302 CTATGATCTATGATATAGCCCTTCTGAAGATGGCTGGAACCTTCCAATTTTGACAGTTTG 361
QY 485 TGGGGCCATATGCTTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTGTATTTTGTACAA 544
Db 362 TGAGGCCCTGTGTCTTTCCAGAGCCAGGAGAACACTTAAATGTGGATTTTATTTGTACAA 421
QY 545 CTGAGGCTGGGGCCGCTTAACTGAAGTGGCGTCTCTCACAAAGTCTTGCAGGAAGTGA 604
Db 422 CTGAGGCTGGGGTCCGCTTGTCCGAAGTGCAGGCTCCACAAAGTCTTGCAGCAGGTAA 481
QY 605 ATCTGCTATTTTACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGGCCCA 664
Db 482 ACCTGCCAATTTTGAATCTCAGGAGGAATGTAGAGCGCTTCTGTGACCTTAAAGAATCCCA 541
QY 665 TCAGTGGGAAGACCTTTCTTTGCACAGGTTTCTCTGATGGAGGAGAGACCATGTGCTAG 724
Db 542 TTACTGGGAAGACCTTTCTCTGCACAGGATCCCTGTATGGAGGAGAGATGCTTTGTCAAG 601
QY 725 GAGATTGAGGAGGTTTCACTCATGTGCGGGAATAAGAAAGGGCGCTGAGCTCTGCTGGTGTG 784
Db 602 GAGATTCAAGAGGGGACACTTATGTGCCCAATAGAAAAGGGCCCTGACACTTTGCTGGAG 661
QY 785 TGACTTC 791
Db 662 TGACTTC 668
RESULT 2
A2319576/c
LOCUS
DEFINITION
1M0039E20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
Clone UUGC1M0039E20 F, DNA sequence.
ACCESSION
A2319576
VERSION
A2319576.1 GI:10370488
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 682)
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0039 row: E column: 20
Seq primer: CGTTGTAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 682.

```


HTC; CAP trapper.
Mus musculus (strain:C57BL/6J) adult female placenta cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1600027G01.

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata,K., Itoh,M., Aizawa,K., Kigaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Natsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staib,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Balarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bernaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ringo,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

5 (bases 1 to 2246)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
Kurthara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
Saito,H., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,D.,
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence[5',
GAGAGAGAGAGCGCGCAATTAATTCGAGTAAATAATTAATTCGCCCCCCC 3'], cDNA
was cleaved with XhoI and SstI. Cloning sites, 5' end:XhoI, 3' end:
SstI. Host: SOLR.
Location/Qualifiers
1. .2246
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:1600027G01"
/db_xref="MGD:MGI:1904755"
/db_xref="taxon:10090"
/clone="1600027G01"
/sex="female"
/tissue_type="placenta"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
94. .1968
/note="data source:SPTR, source key:P03951, evidence:ISS
homolog to COAGULATION FACTOR XI PRECURSOR (EC 3.4.21.27)
(PLASMA THROMBOPLASTIN ANTECEDENT) (PTA)
putative"
/codon_start=1
/protein_id="BAB24114.1"
/db_xref="GI:12838178"
/db_xref="MGD:MGI:1919281"
/translation="MTSLHQLVYFIFPAVSSECVTKFKDIFSQGGDLSTVFPSTP
YCRVLCTHPRCLLFTWASSSDPTKFWACILKDSVTEILPMVNTAGISGYSFK
QSQLSTCKSDVYVNDKGMNYSVVKVARECOERCTDHAHCQETVATGVPFSVD
HRKCLLYKTRTGPTTITKLNGVVSFSLKSGCLNLACIRIDIFNTVLADNLDSV
VAPAFVCRRTCTHPTCTFFTPFSQWPKRESRHLLKLTSSGLSPTRTKSHALS
GFSLQHCRRVPEVFEHPSFYNDTDFLGEELDIDVVKQETCKQTNNAACQKFTFY
SHRLENNRRRCYGLKSSNGSPTRILHGRGGTSGVSLRCKMDNVCTTKINPRVVG
GAASVHGEPMQVTLHTLSQHLGGSGTIIIGNOWITAAHCFSGTETPKLRLVGIQNV
SEINEGFAFFVPEMIITHDOYTAESGYDIALLLKLESAMNVTDFORPCLPSKGDNA
VHTCEGTWGYTALRGVEVSTLQAKVPLVSEECOTRYRRHKINKMKICAGYKKEGG
KDTCKDGGSGPLSKYNGVHVLGVTISWEGCGCKERPVTYVNAKYVDVILETKTQY
"

BASE COUNT 666 a 509 c 536 g 535 t

ORIGIN

Query Match 8.7%; Score 80.4; DB 11; Length 2246;
Best Local Similarity 48.5%; Pred. No. 6.4e-13;
Matches 321; Conservative 0; Mismatches 326; Indels 15; Gaps 3;

QY 150 TAACATTTTTCAGTCGCATCTTTGGAGGAAGCAAGTCGAGAGGGTTCTATCCCTCGCA 209
||||| ||||| ||||| |||||

Db 1245 TAAATCAACCCCGAGAGGAGTGGTGGAGGAGCTGCGTCTTCACGGTGAGTGGCCATGGCA 1304
||||| ||||| ||||| |||||

QY 210 GGTATCTCTGAACAAAGGAGCAAGCATATTGTTGGAGGAAGCATCGTCTTCAACACAGTG 269
||| ||||| ||||| ||||| |||||

Db 1305 GTGACTCTGCACATCAGCCAGGACACTGTGTGGAGGCTTCATCATTTGAACACCATG 1364
||| ||||| ||||| ||||| |||||

QY 270 GTGTATCAGCGGGCTCACTGCATTTGCAACAGAAACATTTGTCTACTTTGAATGTTAC 329
||| ||||| ||||| ||||| |||||

Db 1365 GATACCTGACAGCAGCTCATTTGTTCTTGGATAGAGACACTAAAAGCTCGGTGCTA 1424
||| ||||| ||||| ||||| |||||

Query Match	8.3%	Score 76.6;	DB 9;	Length 596;
Best Local Similarity	51.1%	Pred. No. 3.6e-12;		

Best Local Similarity	51.19	Pred. No.	3.6e-12	
Matches	206	Conservative	0	Mismatches 194
				Indels 3
				Gaps 1

Db 166 TTTCAGTAGGATGTTGGAGGAACCGACACTAGGCAAGGAGCATGGCCTTGGCAGATCA 225

[illegible]

335 GAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAC^TCTCAGCTATTGAAACTGTCA 394

Qy 395 TCATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTGAAGA 454

QY 455 TGGCTGGAGGCTTCCCAATTTTGGCCCACTTTGTGGGGGCCCATATGCTTCCAGAGCTGCGGG 514

Db 463 TGTCTAGCCCATTAATAATACCGAATACATCTCCCAATATGCACTGCCGACTGCTACGG 522

Db 523 CCACATTCCCCCTGGGACTCAATGCTGGTAAACAGGATGGG 565

RESULT 6	BM104667	675 bp	mRNA	linear	EST 21-NOV-2001
LOCUS	BM104667				
DEFINITION	f49hl1.y1 Sugano SJD adult male Danio rerio cDNA clone 5411972.5' similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ; mRNA sequence.				

REFERENCE
AUTHORS
1 (bases 1 to 675)
; Cyprinidae; Danio.
Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 675)

TITLE	JOURNAL	COMMENT
WASNU ZEPHALINA EST PROJECT 1998	Unpublished (1998)	Other ESTs: fv49hll.x1

Email: zbrafish@watson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution information can be found through the I.M.A.G.E.

FEATURES	source
Location/Qualifiers	
1. .596	
/organism="Silurana tropicalis"	
/db_xref="taxon:8364"	
/clone="TCa5012h03"	
/clone_lib="Xgc-gastrula"	
/dev_stage="gastrula (stages 10.5-13 mixed)"	
/lab_host="Escherichia coli XLI-blue"	
/note="Vector: pCSI07; Site_1: EcoRI; Site_2: NotI; cDNA"	

Custom primers for sequencing: 5' end primer
CTTCTCTCTAAAGTCGG and 3' end primer
CGACCTGCAGCTCGAGCACA."

BASE COUNT 169 a 139 c 155 g 179 t
ORIGIN

Query Match 8.1%; Score 74.8; DB 14; Length 642;
Best Local Similarity 51.2%; Pred. No. 1.4e-11;
Matches 175; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 160 AGTCGCATCTTGGAGAGCCCAAGTGGAGAGGGTTCCTATCCCTGGCAGGATCTCTG 219
Db 196 AGCCGGATTATTGGAGAGAGGCGCTGGGCACATCTGTTGGCCATGGCAGGTTTCTG 255
QY 220 AAACAAAGGAGAGAGCATATTGTTGGAGGAGCATCTCTCACACAGTGGGTGATCAGC 279
Db 256 CAGTACATGATGCGCAACCTGTGGTGGCCATCTCCAGCAGCGGTGGTTATATACC 315
QY 280 GCGGCTCAGTCATGCAACAGACAAACATTTGTCTACTTTTGAATGTTACTGCTGGAGAG 339
Db 316 GCGGCTCAGTCGTTTAAACGTTTATAAAAGCCATCTATGTGGAAGCAGTGGTTGGATTG 375
QY 340 TATGACTTATAGCCAGACAGACAGCCAGGAGAGCAACTCTCACTATTGAACTGTCAATCATA 399
Db 376 CACAATTGATATGCGAATGAATCCTCAGCGGAATCCATACAGGTGCAAAAATCTTC 435
QY 400 CATCCACATTTCTCCACCAAGAACCAATGAGTATGATATGCTTCCTTTTGAAGATGGCT 459
Db 436 TCTCACAAGAACTACACAGACAGACAAACGAGAATGATATTCCTTCGCTGAACATGCAG 495
QY 460 GGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTT 501
Db 496 AGTCCGCTGGTTCAGTAAATTTGTAGGCGCTATTGGCGCTT 537

RESULT 10

BQ258513 642 bp mRNA linear EST 06-MAY-2002
LOCUS f282f02.y1 Sugano SJD adult male Danio rerio cDNA clone 5914899 5'
DEFINITION similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;
mRNA sequence.

ACCESSION BQ258513.1 GI:20459272

VERSION EST.

KEYWORDS zebrafish.

SOURCE Danio rerio

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

REFERENCE 1 (bases 1 to 642)

AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
and Willson, R.

TITLE WashU Zebrafish EST Project 1998

JOURNAL Unpublished (1998)

COMMENT Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@wustl.edu

Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA

Sequencing by: Washington University Genome Sequencing Center Clome

distribution information can be found through the I.M.A.G.E.

Consortium/LLNL, send email to: info@image.llnl.gov

Seq primer: T3 ET from Amsersham

High quality sequence stop: 526.

FEATURES

source

1..642 Location/Qualifiers

/organism="Danio rerio"

/db_xref="taxon:7955"
/clone="5914899"
/clone_lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="Vector: pME18S-FL3; Site 1: DraIII (CACCATGTG);
Site 2: DraIII (CACTGTGTG); 1st strand cDNA was primed
with an oligo(dT) primer (ATGTGCGCTTTTTTTTTTTTTTTT);
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGGCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTCTCTAAAGTCGG and 3' end primer
CGACCTGCAGCTCGAGCACA."

BASE COUNT 170 a 138 c 155 g 179 t
ORIGIN

Query Match 8.1%; Score 74.8; DB 14; Length 642;
Best Local Similarity 51.2%; Pred. No. 1.4e-11;
Matches 175; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 160 AGTCGCATCTTGGAGAGCCCAAGTGGAGAGGGTTCCTATCCCTGGCAGGATCTCTG 219

Db 199 AGCCGGATTATTGGAGAGAGGCGCTGGGCACATCTGTTGGCCATGGCAGGTTTCTG 258

QY 220 AAACAAAGGAGAGAGCATATTGTTGGAGGAGCATCTCTCACACAGTGGGTGATCAGC 279

Db 259 CAGTACATGATGCGCAACCTGTGGTGGCCATCTCCAGCAGCGGTGGTTATATACC 318

QY 280 GCGGCTCAGTCATGCAACAGACAAACATTTGTCTACTTTTGAATGTTACTGCTGGAGAG 339

Db 319 GCGGCTCAGTCGTTTAAACGTTTATAAAAGCCATCTATGTGGAACGACGTGGTTGGATTG 378

QY 340 TATGACTTAAAGCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAAACGTGCATCATA 399

Db 379 CACAATTGGATATGGAATGAATCTCTCAGGGAATCCATACAGGTGCAAAAATCTTC 438

QY 400 CATCCACATTTCTCCACCAAGAACCAATGGACTATGATATTCCTTTTGAAGATGGCT 459

Db 439 TCTCACAAGAACTACAACAGACAGACAAACGAGAATGATATTCCTCTGCTGAAATGCAG 498

QY 460 GGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTT 501

Db 499 AGTCCGCTGGTGTTCAGTAAATTTGTAGGCGCTATTGGCGCTT 540

RESULT 11

BMI83510

LOCUS

DEFINITION

BM183510

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

1 (bases 1 to 645)

AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,

Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,

Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.

and Willson, R.

similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;

mRNA sequence.

645 bp

linear

EST 11-DEC-2001

fv64d10.y1 Sugano SJD adult male Danio rerio cDNA clone 5413291 5'

similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;

mRNA sequence.

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Socioeconomic Status on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Education	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Training on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Socioeconomic Status on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clome
distribution information can be found through the I.M.A.G.E.
Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T3 ET from Amersham
High quality sequence stop: 514.

FEATURES

```

Location/Vadarrilais
1. .645
/organism="Danio rerio"
/db_xref="taxon:7955"
/cclone="5413291"
/cclone.lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="vector: pME18S-FL3; Site 1:
Site 2: DraIII (CACATGCTGT); 1st strand
with an oligo(dT) primer [AGTGCGC]
double-stranded cDNA was ligated to
[TTGTGGCCACTACGT], digested and cloned
sites of the pME18S-FL3 vector (5'
CACCATCTGT). XhoI should be used to
insert. Size selection was performed
<1.5kb. Library constructed and deposited
Sugano (University of Tokyo Institute
Custom primers for sequencing: 5'
CTTCCTGCCTAAAGATGGTC and 3' end primer
CGACCTTCGCTGACGACCA."

```

BASE COUNT	168 a	139 c	155 g	183 t
ORIGIN				

Query Match 8.1%; Score 74.8; DB 13; Length 645;
Best Local Similarity 51.2%; Pred. NO. 1.4e-11;
Matches 175; Conservative 0; Mismatches 167; Indels 0

160	Qy	AGTCGCATTC	TGTGGAGGAC	CCAAAGTC	GGAGGGTT	CCCTAT	CCCTGGCAGG	TATCTGT	219
202	Db	AGCCGGAT	TATGTGGAAAG	GAGGCC	TGGGCACAT	TCTGGGCCAT	GCGAGGTT	CTCTGT	261
220	Qy	AAACAAGG	CAGACATAT	TTGTGGAG	GAAGCATC	GTCACCA	CACAGATGGG	TGATCACG	279
262	Db	CAGTACAA	TGATGTGCCAA	CCGTGTGG	TGGTGC	CATCTCG	CACACGCG	TGGTGTATTACC	321
280	Qy	CGCGCT	CACATGC	ATTGC	AAACAGAA	CATTGTGT	CTACTTTG	ATGTTACT	339
322	Db	GCCGGT	CAC	TGCTTTAA	ACGTTAT	AAAGCC	ATCTATGT	GGAACG	381
340	Qy	TATGACT	TAAGCC	ACAGAC	CCACGAG	AGACAA	ACTCTC	ACTATTG	399
382	Db	CACAAT	TTGGAT	TATGCA	TGATCTC	ACGGGA	ATCCATAC	AGGTGCA	441
400	Qy	CATCCAT	TTCTCC	ACCAAGAA	ACCAATGG	ACATAT	TG	CCCTTTG	459
442	Db	TCTCACA	AGAACTACA	CCACG	AGACACA	ACGAGA	ATGATTTG	CCCTGCTG	501
460	Qy	GGAGCCT	TCCAA	TTGGCC	ACTTTGTGGG	CCCATATG	CTTT		501
502	Db	AGTCCG	CTGGTGT	TTCAGT	TAATTTG	CAGG	CCTATTG	CGC	543

RESULT. 12

RESOLUTION	LOCUS	BM102995	652 bp	mRNA	linear	EST
12	BM102995	LOCUS	652 bp	mRNA	linear	21-NOV-2001
		DEFINITION	fv39f04.v1	Sucano SJD adult male	Danio rerio cDNA clone 5411070.5	

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

ORGANISM

REFERENCE AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1981, Vol. 5, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Psychology	1982, Vol. 74, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1983, Vol. 86, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1984, Vol. 87, No. 5, pp. 41-50
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1985, Vol. 88, No. 6, pp. 51-60
7. The Importance of Teacher Professionalism	Journal of Educational Research	1986, Vol. 89, No. 7, pp. 61-70
8. The Effect of Teacher Salary on Student Achievement	Journal of Educational Research	1987, Vol. 90, No. 8, pp. 71-80
9. The Role of the School in the Community	Journal of Educational Research	1988, Vol. 91, No. 9, pp. 81-90
10. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1989, Vol. 92, No. 10, pp. 91-100

FEATURES
SOURCE

	BASE COUNT	ORGTN
0	100	0
1	90	1
2	80	2
3	70	3
4	60	4
5	50	5
6	40	6
7	30	7
8	20	8
9	10	9
10	0	10

Query Match	8.1%	Score 74.8;	DB 13;	Length 652;
Best Local Similarity	51.2%;	Pred. No. 1.4e-11;		
Matches 175;	Conservative	0;	Mismatches 167;	Indels 0;
				Gaps 0;

Qy	160	AGTCGACTCTTGGAGGAGCCCAAGTGGAGAAAGGGTTCCCTATCCCTGGCAGGTTATCTCTG	219
Db	194	AGCCGGATTAATTGGTGGAAAGAGGCCCTGGGCACATTCGTGGCCCATGGCAGGTTCTCTG	253
Qy	220	AAACAAGGCGAAGACATATTTTGTGGAGGAAGCATCGTCTCACCACAGTGGGGTGCATCAGC	279
Db	254	CAGTACAATGATGTGCCNAACCTGTGGTGGTGCCCATCTCGACACGGGTGGGTTATTACC	313
Qy	280	CGGGCTCACTGCATTGCAACAGAAACATCTGTCTACTTTTGAATGTTACTGCTGGAGAG	339
Db	314	CGCGTCACTGCTTTAAACGTTATAAAAGCCATCTATGTGGAACGCAGTGGTTGGATTG	373
Qy	340	TATGACTTAACCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTCATCATYA	399
Db	374	CACAATTTGGATATATGCGAATGNAATCCTCACGGGAATCCATACAGAGTGCAAAAAATCTTC	433
Qy	400	CATCCACATTTTCTCCACCACGAACCAGTGGACTATGATATATGCGCTTTTTCGAAGATGGCT	459

similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;
mRNA sequence.
BM102995
BM102995.1 GI:17034065
EST.
zebrafish.
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Neopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes
; Cyprinidae; Danio.
Phases 1 to 652)

Clark, M., Johnson, S. L., Lebrach, H., Lee, R., Li, F., Marra, M., Eddy, J. (bases 1 to 652)

WashU Zebrafish EST Project 1998
 Unpublished (1998)
 Contact: Stephen L. Johnson

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
High quality sequence stop: 527.

FEATURES
SOURCE

```

1. .652
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="5411070"
/clone_lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/notes="Vector: pME18S-FL3; Site.1: DraIII (CACCATGTG);
Site.2: DraIII (CAGCTGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGGCTTACGTG], digested and cloned into distinct draIII
sites of the pME18S-FL3 vector (5' site CAGCTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAGCTGGC and 3' end primer
CGACCTGCGAGCTCGAGCA."

```

CGACCTGCAGCTCGAGCACA.		
172 a	142 c	156 q 182 t

9
0
0
4

n

0
0
-
4

6
4
-
4

Query Match	8.1%	Score 74.8;	DB 13;	Length 652;
Best Local Similarity	51.2%;	Pred. No. 1.4e-11;		
Matches 175;	Conservative	0;	Mismatches 167;	Indels 0;
				Gaps 0;

Qy	160	AGTCGACTCTTGGAGGAGCCCAAGTGGAGAAGGGTTCCCTATCCCTGGCAGGTAATCTCTG	219
Db	194	AGCCGGATTAATTGGTGGAAAGAGGCCCTGGGCACATTCGTGGCCCATGGCAGGTTCTCTG	253
Qy	220	AAACAAGGCGAAGACATATTTTGTGGAGGAAGCATCGTCTCACCACAGTGGGGTCATCAGC	279
Db	254	CAGTACAATGATGTGCCNAACCTGTGGTGGTGCCATCTCGACACGGGTGGGTTATTACC	313
Qy	280	CGGGCTCACTGCATTGCAACAGAAACATCTGTCTACTTTTGAATGTTACTGCTGGAGAG	339
Db	314	GCCGTCACCTGCTTTAAACGTTATAAAAGCCATCTATGTGGAACGCAGTGGTTGGATTG	373
Qy	340	TATGACTTAACCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTCATCATYA	399
Db	374	CACAATTTGGATATATGCGAATGNAATCCTCACGGGAATCCATACAGAGTGCAAAAATCTTC	433
Qy	400	CATCCACATTTTCTCCACCAGAAACCAGTGGACTATGATATATGCGCTTTTTCGAAGATGGCT	459

double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed and donated by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACAC.

BASE COUNT	174 a	142 c	158 g	182 t
ORIGIN				
Query Match	8.1%;	Score 74.8;	DB 13;	Length 656;
Best Local Similarity	51.2%;	Pred. No. 1.4e-11;		
Matches 175;	Conservative 0;	Mismatches 167;	Indels 0;	Gaps 0;
QY 160	AGTCGCATTCTTGGAGGAGCCAAAGTGGAGAGGGTTCTCTATCCCTGGGAGGATATCTCTG	219		
Db 195	AGCCGGATTATTGGTGGAAAGGAGGCTGGGCACATTCGTGGCCATGGCGAGGTTTCTCTG	254		
QY 220	AAACAAGGCAGAGCATATTTGTGGAGGAGCATGCTCTCACCACAGTGGGTGATCAGC	279		
Db 255	CAGTACAATGATGTGCCAACCTGTGGTGGTGGCCATCTCGACAGCGGTGGGTATATACC	314		
QY 280	GGGGCTCACTGCTCAACACAGAAACATTTGTCTACTTTGAATTTACTGTCTGTGAGAG	339		
Db 315	GCCGGTCACTGCTTTAAAGCTTATAAAAGCCATCTATGTGNACCCAGTGGTGTGATTTG	374		
QY 340	TATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATTGAAACTGTCAATCATYA	399		
Db 375	CACAATTTGGATAATCGAATGAATCCTCACGGGAATCCATACAGTGCAGAAATCTTC	434		
QY 400	CATCCACATTTCTCCACCAAGAACCAATGGACTATGATATGCCCTTTTGAAGATGGCT	459		
Db 435	TCTCAAGAAGACTACACACAGAGAACAGAGAGATGATTTGCCCTGCTGAAACTGCAG	494		
QY 460	GGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCCATATGTCTT	501		
Db 495	AGTCCCGCTGTTTCAGTAATTTGTGAGGCCCTATTGGCGTT	536		

RESULT 15				
LOCUS	BQ075190	657 bp	mRNA	linear
DEFINITION	fya3a08.y2 Sugano SJD adult male Danio rerio cDNA clone 5602455.5, similar to SWTMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ; mRNA sequence.			
ACCESSION	BQ075190			
VERSION	BQ075190.1	GI:19904227		
KEYWORDS	EST.			
SOURCE	zebrafish.			
ORGANISM	Danio rerio			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.			
AUTHORS	1 (bases 1 to 657) Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.			
TITLE	WashU Zebrafish EST Project 1998			
JOURNAL	Unpublished (1998)			
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbrafish@wustl.edu Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA Sequencing by: Washington University Genome Sequencing Center Clone			

distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T3 ET from Amersham
High quality sequence stop: 522.

FEATURES

location/Qualifiers	
1..857	
/organism="Danio rerio"	
/db_xref="taxon:7955"	
/clone_lib="5602455"	
/clone_lib="Sugano SJD adult male"	
/sex="male"	
/tissue_type="whole body"	
/dev_stage="adult"	
/lab_host="DH10B (phage resistant)"	
/note="Vector: pME18S-FL3; Site.1: DraIII (CACCATGTG); Site.2: DraIII (CACTGTGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed and donated by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACAC."	
171 a	141 c 159 g 186 t
BASE COUNT	
ORIGIN	

Search completed: March 25, 2003, 06:30:34
Job time : 1860 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 25, 2003, 06:47:42 ; Search time 16 Seconds

(without alignments)
3387.314 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 1668

Sequence: 1 atgagtcctcaaatgcttat.....acatccaaactgtaactaa 921

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09735713/runat_18032003_124355_23161/app_query.fasta_1.1095
-DB=issued_patents_AA -QFMT=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713@cgn1.13 @runat_18032003_124355_23161 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	467	28.0	454	4	US-09-518-046-2
2	460	27.6	248	4	US-08-944-483-63
3	446.5	26.8	455	4	US-09-261-416-2
4	446	26.7	638	2	US-08-681-151-3
5	442.5	26.5	798	1	US-08-200-900A-2
6	442.5	26.5	798	5	PCT-US94-00616-2
7	440	26.4	416	2	US-09-000-846-2
8	439.5	26.3	328	4	US-09-386-642-11
9	437.5	26.2	256	2	US-09-027-337-3
10	437.5	26.2	256	4	US-09-644-600-3
11	436	26.1	319	4	US-09-386-642-12
12	432.5	25.9	255	4	US-08-944-483-67

13	432.5	25.9	400	4	US-09-004-731-30	Sequence 30, Appl
14	432.5	25.9	400	4	US-09-004-731-33	Sequence 33, Appl
15	432.5	25.9	400	4	US-08-749-699-30	Sequence 30, Appl
16	432.5	25.9	400	4	US-08-749-699-33	Sequence 33, Appl
17	432.5	25.9	400	4	US-09-004-729-30	Sequence 30, Appl
18	432.5	25.9	400	4	US-09-004-729-33	Sequence 33, Appl
19	430.5	25.8	387	4	US-09-032-215-8	Sequence 8, Appl
20	430.5	25.8	387	4	US-09-032-215-13	Sequence 13, Appl
21	430	25.8	299	4	US-08-944-483-66	Sequence 66, Appl
22	426	25.5	242	4	US-09-004-731-36	Sequence 36, Appl
23	426	25.5	242	4	US-08-749-699-36	Sequence 36, Appl
24	426	25.5	242	4	US-09-004-729-36	Sequence 36, Appl
25	425.5	25.5	235	3	US-08-807-151-3	Sequence 3, Appl
26	425.5	25.5	235	4	US-09-478-957-3	Sequence 3, Appl
27	425.5	25.5	290	4	US-09-386-653A-7	Sequence 7, Appl
28	425	25.5	315	4	US-09-386-653A-9	Sequence 9, Appl
29	421	25.2	317	4	US-09-386-629-7	Sequence 7, Appl
30	419.5	25.1	235	4	US-08-944-483-65	Sequence 65, Appl
31	419.5	25.1	237	4	US-08-163-919A-3	Sequence 3, Appl
32	419.5	25.1	237	5	PCT-US94-14073-3	Sequence 3, Appl
33	414	24.8	435	4	US-09-008-271A-6	Sequence 6, Appl
34	413	24.8	327	4	US-09-386-629-8	Sequence 8, Appl
35	413	24.8	423	4	US-09-656-002-2	Sequence 2, Appl
36	408	24.4	238	4	US-08-944-483-64	Sequence 64, Appl
37	407	24.4	283	3	US-08-807-151-1	Sequence 1, Appl
38	407	24.4	283	4	US-09-478-957-1	Sequence 1, Appl
39	406	24.3	254	4	US-09-578-303-5	Sequence 5, Appl
40	406	24.3	492	4	US-09-342-749-2	Sequence 2, Appl
41	406	24.3	492	4	US-09-691-840-2	Sequence 2, Appl
42	405	24.3	231	2	US-09-027-337-6	Sequence 6, Appl
43	405	24.3	231	4	US-09-644-600-6	Sequence 6, Appl
44	403	24.2	314	4	US-09-008-271A-3	Sequence 3, Appl
45	402	24.1	267	2	US-09-016-366A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-09-518-046-2

; Sequence 2, Application US/09518046
; Patent No. 6294663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof
; FILE REFERENCE: D6192CIP
; CURRENT APPLICATION NUMBER: US/09/518,046
; CURRENT FILING DATE: 2000-03-02
; EARLIER APPLICATION NUMBER: 09/261,416
; EARLIER FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 2
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: complete amino acid sequence of TADG-12
; OTHER INFORMATION: protein
US-09-518-046-2

Alignment Scores:
Pred. No.: 7,48e-43 Length: 454
Score: 467.00 Matches: 99
Percent Similarity: 51.95% Conservative: 61
Best Local Similarity: 32.11% Mismatches: 114
Query Match: 28.00% Indels: 34
DB: 4 Gaps: 9

US-09-735-713A-1 (1-921) x US-09-518-046-2 (1-454)

QY 1 ATGAGTCTCAAAATGCTTTATAGCAGGACAGCTGATTTTACTAGGAATAGTCTTT 60
:::||||::: |||::: |||
:::||||::: |||::: |||

Db 171 ValSerIleAspHisLeuLeuProAspAspLysValThrAlaLeuHisHisSerValTyr 190
Qy 61 TTTGAACRAGT---AAATCTGCARCTCTTTTCGCTCCCAAGTCCAGTGTGGCAG 117
Db 191 ValArgGluGlyCysAlaSerGlyHisValValThrLeuGlnCysThrAlaCysGlyHis 210
Qy 118 AGTCTGTTAAGGTACAGCCTTGAATATTATTTAACTATTTTCAGTCGCTATCTTGGAGGA 177
Db 211 ArgArg-----GlyTyrSerSerArgIleValGlyGly 221
Qy 178 ACCCAAGTGGAGAGGTTCTATCCCTCGGAGGTATCTCTGAACAAAGGACGAAGCAT 237
Db 222 AsnMetSerLeuLeuSerGlnTyrProTrpGlnAlaSerLeuGlnPheGlnGlyTyrHis 241
Qy 238 ATTTGTGGAGGAGCATGCTCACCACAGTGGGTGATCAGCGGCTCACTGCTATTCGA 297
Db 242 LeuGlyGlySerValIleThrProLeuTrpIleThrAlaAlaHisCysValTyr 261
Qy 298 AACAGAAACATTTGCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACA 357
Db 262 AspLeuTyrLeuProLysSerTrpThrIleGlnValGlyLeuValSerLeuLeuAsp--- 280
Qy 358 GACCCAGGAGACAACTCCTATTTGAAGTGGTGGTCTCATCATCATCATCATCATCAT 417
Db 281 AsnProAlaProSerHisLeu---ValGluLysIleValTyrHisSerLysTyrLysPro 299
Qy 418 AAGAAACCAATGGACTATGATATTGCTTTTGAAGATGGTGGAGCTTCCAATTTGGC 477
Db 300 LysArg---LeuGlyAsnAspIleAlaLeuMetLysLeuAlaGlyProLeuThrPheAsn 318
Qy 478 CACTTTGTGGGGCCATATGCTTCCAGAGCTGGCGGAGCAATTTAGGCTGGTTTATT 537
Db 319 GluMetIleGlnProValCysLeuProAsnSerGluGluAsnPheProAspGlyLysVal 338
Qy 538 TGTACACTGACGCTGGGCGGCTTAACTGAAGTGGCTCTCTCACAAGTCTTCGAG 597
Db 339 CysTrpThrSerGlyTrpGlyAlaThrGluAspGlyGlyAspAlaSerProValLeuAsn 358
Qy 598 GAAGTGAATCTGCCTATTTTGAACCTGGAGAGTGTGTGGCAGCTCTGTAAACACTAAAG 657
Db 359 HisAlaAlaValProLeuIleSerAsnLysIleCys-----AsnHis 372
Qy 658 AGGCCCATCAGTGG-----AAGACCTTTCTTGGCAGAGTTTTCGTGATGGA 705
Db 373 ArgAspValTyrGlyGlyIleIleSerProSerMetLeuCysAlaGlyTyrLeuThrGly 392
Qy 706 GGGAGAGACGATGTCAGGAGATTCAGGAGTTCATCATGTCCCGGAATAAGAAAGG 765
Db 393 GlyValAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGluArg--- 411
Qy 766 GCCTGGACTCTGGCTGTGCTACTCTCTGGGTTTGGGCTTGTGCTGAGGCTGGAGAAAC 825
Db 412 LeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys----- 425
Qy 826 AATGTGAGAAAGTATCAAGGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTT 885
Db 426 -----AlaGluValAsnLysProGlyValTyrThrArgValThrSerPheLeu 441
Qy 886 TCCTGGATCCACGAACATCCAA 909
Db 442 AspTrpIleHisGluGlnMetGlu 449

RESULT 2

US-08-944-483-63
; Sequence 63, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-63

Alignment Scores:

Pred. No.:	3,4e-42	Length:	248
Score:	460.00	Matches:	90
Percent Similarity:	55.12%	Conservative:	50
Best Local Similarity:	35.43%	Mismatches:	96
Query Match:	27.58%	Indels:	18
DB:	4	Gaps:	5

US-09-735-713A-1 (1-921) x US-08-944-483-63 (1-248)

Qy	166	ATCTTGGAGAACCCAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTGAACAA	225
Db	1	IleValGlyGlyThrAsnSerSerTrpGlyGlyTrpGlnValSerLeuGlnVal	20
Qy	226	AGG-----CAGAACGATATTGTGGAGGAGACATCGTCTCCACACAGTGGGTGATC	276
Db	21	LysLeuThrAlaGlnArgHisLeuCysGlyGlySerLeuIleGlyHisGlnTrpValLeu	40
Qy	277	ACGGCGCTCAGTCGATTCGAAACAGAAACATTTGTCTACTTTGAATGTTACTGTGGA	336
Db	41	ThrAlaAlaHisCysPheAspGlyLeuProLeuGlnAspValTrpArgIleTyrSerGly	60
Qy	337	GAGTATGACTTAAGCCAGACAGACAGACCCAGGAGAGAACTCTCAGTATTGAACGTGTCATC	396
Db	61	IleLeuAsnLeuSerAspIleThrLysAspThrProPheSerGlnIleLysGluIlelle	80
Qy	397	ATACATCCAGTATCTCCACCAAGAACCAATGATGATATTCCTTTTGAAGATG	456
Db	81	IleHisGlnAsnTyrLysValSerGluGly---AsnHisAspIleAlaLeuIleLysLeu	99
Qy	457	GCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCCCAATATGTCTTCCAGAGCTCGGGAG	516

```

Db 100 GlnAlaProLeuAsnFyrThrGluPheGlnLysProIleCysLeuProSerLysGlyAsp 119
QY 517 CAATTTGAGGCTGTTTATTTGACAACTGCAGGCTGGCGCGCTTAACCTGAAGTGGC 576
Db 120 ThrSerThrIleFyrThrAsnCysTrpValThrGlyTrpGlyPheSerLysGluLysGly 139
QY 577 GTCCTCTCAACAGTCTTGCAAGAGTGAATCTGCCTATTATTTGACCTGGGAAGAGTGTG 636
Db 140 GluIleGlnAsnIleLeuGlnLysValAsnIleProLeuValThrAsnGluCysGln 159
QY 637 GCAGCTCTGTTAACATAAAGAGGCCATCAGTGGGAAGACCTTTCTTGCACAGGTTT 696
Db 160 LysArgTyrGlnAspFyrLys-----IleThrGlnArgMetValCysAlaGlyTyr 176
QY 697 CCTGATGGAGGAGACACCATCTCAGGAGATTCAGGAGGTTCACCTCATGTCGCCGAAT 756
Db 177 LysGluGlyCysLysAspAlaCysLysGlyAspSerGlyGlyProLeuValCysLysHis 196
QY 757 AAGAAAGGGCGCTGGACTCTGGCTGTGTGACTTCTCGGGGTTTGGGCTGTGTGTCGAGGC 816
Db 197 ---AsnGlyMetTrpArgLeuValGlyIleThrSerTrpGlyGluCysAlaArg--- 214
QY 817 TGGAGAAACAATGTGAGGAAAAAGTATCAAGGATCCCTCGGATCTTCACAGACATTA 876
Db 215 -----ArgGluGlnProGlyValThrLysValAla 225
QY 877 AAGTGCCTTCTCGGATCCAGCAGACATCCAACTGGTAAC 918
Db 226 GluTyrMetAspTrpIleLeuGluLysThrGlnSerSerAsp 239
RESULT 3
US-09-261-416-2
; Sequence 2, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of TADG-12 encoded by nucleotides
; Patent No. 6291663
US-09-261-416-2
Alignment Scores:
Pred. No.: 1,34e-40 Length: 455
Score: 446.50 Matches: 98
Percent Similarity: 49.03% Conservative: 57
Best Local Similarity: 31.01% Mismatches: 112
Query Match: 26.77% Indels: 49
DB: 4 Gaps: 10
US-09-735-713a-1 (1-921) x US-09-261-416-2 (1-455)
QY 1 ATGAGTCTCAAAATGTTATAGCAGGAACAAGCTGATTTTACTAGTAATAGTCTTT 60
Db 171 ValSerIleAspHisLeuLeuProAspLysValThrAlaLeuHisSerValTyr 190
QY 61 TTTGAACRAGGT---AAATCTGCARTCTTTCGCTCCCAAGCTCCCGAGTGTGGCGAG 117
Db 191 ValArgGluGlyCysAlaSerGlyHisValValThrLeuGlnCysThrAlaCysGlyHis 210
QY 118 AGTCTGTTAAGGTACAGCCTTGGAAATTTTAAACATTTTCAGTCGCATCTCTTGGAGGA 177

```

```

Db 211 ArgArg-----GlyTyrSerSerArgIleValGlyGly 221
QY 178 AGCCRAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAACAAAGGAGCAAT 237
Db 222 AsnMetSerLeuLeuSerGlnTrpProTrpGlnAlaSerLeuGlnPheGlnGlyTyrHis 241
QY 238 ATTTGTGGAGGAGCATCGTCTCACCAAGTGGGTATCAGCGCGCTCACATGCAATGCA 297
Db 242 LeuCysGlyGlySerValIleThrProLeuTyrIleThrAlaAlaHisCysValTyr 261
QY 298 ACAGAAACATTTGTCTACTTTGAATGTACTGCTGGAGAGTATGACTTAAAGCCAGACA 357
Db 262 AspLeuTyrLeuProLysSerTrpThrIleGlnValCysLeuValSerLeuLeuAsp--- 280
QY 358 GACCCAGGAGAGCAAACTCTCATAATTGAAGTGTATCATATCATATCATATCTCCACC 417
Db 281 AsnProAlaProSerHisLeu---ValGluLysIleValTyrHisSerLysTyrLysPro 299
QY 418 AAGAAACCAATGACTATGATATTGCTTTCAGAGTGCAGGAGCAATTTGAGGCTGGTTTAT 477
Db 300 LysArg---LeuGlyAsnAspIleAlaLeuMetLysLeuAlaGlyProLeuThrPheAsn 318
QY 478 CACTTTGTGGGCGCCATATGCTTCCAGAGTGCAGGAGCAATTTGAGGCTGGTTTAT 537
Db 319 GluMetIleGlnProValCysLeuProAsnSerGluGluAsnPheProAspGlyLysVal 338
QY 538 TGTACAACTGCAGGCTGGGCGCTTAACCTGAAGTGGGCTCTCTCAACAAGCTTTCAG 597
Db 339 CysTrpThrSerGlyTrpGlyAlaThrGluAspGlyGlyAspAlaSerProValLeuAsn 358
QY 598 GAAGTGAATCTGCTATTGAC-----TGG 624
Db 359 HisAlaAlaValProLeuIleSerAsnLysAspLeuGlnProGlnGlyArgValArgTrp 378
QY 625 GAAGAGTGTGTGGCAGCTCTGTAAACATAAGAGAGGCCATCAGTGGGAGACCTTTCT 684
Db 379 HisHisLeuProLeuHisAlaLeu----- 386
QY 685 TGCACAGGTTTCTCTGATGAGGAGAGCA---GACCGATGTGAGGAGATTCAGGAGTTCA 741
Db 387 ---ArgGlyLeuProAspGlyTrpArgTrpAsnSerCysGlnGlyAspSerGlyGlyPro 405
QY 742 CTCATGTGCGGAATAAGAAAGGGGCTGAGCTCTGCTGGTGTGCTGCTGCGGTTTG 801
Db 406 LeuValCysGlnGluArgArg---LeuTrpLysLeuValGlyAlaThrSerPheGlyIle 424
QY 802 GGCTGTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAAGGATCCCTCGGATC 861
Db 425 GlyCys-----AlaAspValAsnLysProGlyVal 434
QY 862 TTCACAGACATTAAGTAAAGTGTCTTCTGATCCAGCAACATCCAA 909
Db 435 TyrThrArgValThrSerPheLeuAspTrpIleHisGlnMetGlu 450
RESULT 4
US-08-681-151-3
; Sequence 3, Application US/08681151
; Patent No. 5869637
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Braxton, Scott Michael
; APPLICANT: Goli, Surya
; TITLE OF INVENTION: A NOVEL HUMAN KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304

```


US-09-000-846-2

; Sequence 2, Application US/09000846
; Patent No. 5981830
; GENERAL INFORMATION:
; APPLICANT: WU, QINGYU
; APPLICANT: SADLER, JASPER
; TITLE OF INVENTION: KNOCKOUT MICE AND THEIR PROGENY WITH
; TITLE OF INVENTION: DISRUPTED SERINE PROTEASE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,846
; FILING DATE: 30-DEC-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/866,058
; FILING DATE: 30-MAY-1997

; ATTORNEY/AGENT INFORMATION:
; NAME: LEBOWITZ, RICHARD M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: BERLX 65P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 416 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-000-846-2

Alignment Scores:

Pred. No.:	Score:	Length:
109	440.00	416
152	49.47%	99
169	34.98%	41
163	26.38%	99
229	26.38%	44
183	26.38%	10

US-09-735-713A-1 (1-921) x US-09-000-846-2 (1-416)

QY	109	TGTGGCAGAGTCTGTTAAGTACAGCCTTGGAAATTTTAAACATTTTCAGTCGCATT	168
Db	152	CysGlyArgArgLysLeuProVal-----AspArglie	162
QY	169	CTTGAGGAGCAAGTGGAGAGGGTCTATCCCTGGCAGGTATCTGTGAACAAAGG	228
Db	163	ValGlyGlyGlnAspSerLeuGlyArgTrpProTrpGlnValSerLeuArgTrpAsp	182
QY	229	CAGAAGCATATTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGCGCGCTCAC	288
Db	183	GlyThrHisLeuGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAlaAlaHis	202
QY	289	TGCATTGCAACAGAAAC---ATTGTGTACTTTGAATGTTACTCTGGAGAGTATGAC	345
Db	203	CysProGluArgAsnArgValLeuSerArgTrpArgValPheAlaGlyAla-----	220
QY	346	TTAAGCCAGACAGACCCAGGAGGAGCAACTCTCATTGTAACCTGTCATCATCATCA	405
Db	221	ValAlaArgThrSerProHisAlaValGlnLeuGlyValGlnAlaValIleTyrHisGly	240

QY	406	CAATTC-----TCCACCAAGAAACCAATGAGCATATGATATTCGCTTTT	450
Db	241	GlyTyrLeuProPheArgAspProThrIleAspGluAsnSerAsnAspIleAlaLeuVal	260
QY	451	AAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGCTCTCCAGAGCTG	510
Db	261	HisLeuSerSerSerLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla	280
QY	511	CGGAGCAATTTGAGGCTGGTATTTATTGTACAACTCCAGGCTGGGCTTAACTGAA	570
Db	281	GlyGlnAlaLeuValAspGlyLysValCysThrValThrGlyTrpGlyAsnThrGlnPhe	300
QY	571	GGTGGGCTCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCCTATTGACCTGGGAAGAG	630
Db	301	TyrGlyGlnGlnAlaMetValLeuGlnGluAlaArgValProIleIleSerAsnGluVal	320
QY	631	TGTGTGGCAGCTCTGTTAAACATAAAGAGGCC-----ATCAGTGGG	672
Db	321	Cys-----AsnSerProAspPheTyrGlyAsnGlnIleLysPro	333
QY	673	AGACCTTTTTCACAGGTTTTCCTGATGGAGGAGAGACGATGTGAGGAGATTCA	732
Db	334	LysMetPhe---CysAlaGlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSer	352
QY	733	GGAGTTCACTCATGTGCCGGAATAAG-----AAAGGGCCTGGACTCTGGCTGGT	783
Db	353	GlyGlyProPheValCysGluAspSerIleSerGlyThrSerArgTrpArgLeuCysGly	372
QY	784	GTGACTTCTGGGTTTGGCTGTGCTCGAGGCTGGAGAAACAATGTGAGGAAAGTGT	843
Db	373	IleValSerTrpGlyThrGlyCysAlaLeu-----AlaArgLys-----	385
QY	844	CAAGGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTTCCCTGGATCCACGAAAC	903
Db	386	-----ProGlyValTyrThrLysValThrAspPheArgGluTrpIlePheLysAla	402
QY	904	ATCCAAACT 912	
Db	403	IleLysThr 405	

RESULT 8

US-09-386-642-11
; Sequence 11, Application US/09386642

; Patent No. 6420157

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew

; APPLICANT: Qi, Jenson

; APPLICANT: Andrade-Gordon, Patricia

; TITLE OF INVENTION: Zymogen Activation System

; FILE REFERENCE: ORT-1028

; CURRENT APPLICATION NUMBER: US/09/386,642

; CURRENT FILING DATE: 1999-08-31

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 11

; LENGTH: 328

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene

; OTHER INFORMATION: with homo sapien serine protease catalytic domain

US-09-386-642-11

Alignment Scores:

Pred. No.:	Score:	Length:
6	439.50	328
48.71%	52	
31.94%	120	
26.35%	39	
4	8	

US-09-735-713A-1 (1-921) x US-09-386-642-11 (1-328)


```
QY 7 CTCAATGCTTATAAGCAGGAAAGCTGATTTTACTAGCAATAGTCTTT----- 60
Db 12 LeuLeuLeuLeuValValSerAsnLeuLeuLeuLeuValValSerAspTyr 31
QY 61 -----TTTGAACRAGGTAAATCTGCARTCTTTTCGCTCCCAAGCTCCAGTTGGG 114
Db 32 LysAspAspAspValAspAlaAlaLeuAlaAlaPro----- 45
QY 115 CAGAGTGTGTTAAGCTACAGCTTGGAAATATTATTTAAACATTTTCAGTCGCATTTGGA 174
Db 46 -----PheAspAspAspLysIleValGly 54
QY 175 GGAAGCAAGTGGAGAGGTTCTCTATCCCTGGCAGTATCTCTGAACAAAGGCAGAAG 234
Db 55 GlyTyrAlaLeuLeuAlaGlyGlnTyrProTyrGlnValSerIleThrTyrGluGlyVal 74
QY 235 CATATTTGGAGGAGCATCTCTCACACAGTGGGTGATCAGCGGCTCACTGCATT 294
Db 75 HisValCysGlySerLeuValSerGluGlnTyrValLeuSerAlaAlaHisCysPhe 94
QY 295 GCAACAGAAACATCTGTCTACTTTGAATGTTACTGCTGGAGATGACTTTAACCCAG 354
Db 95 ProSerGluHisLysGluAlaTyrGluValLysLeuGlyAlaHisGlnLeuAspSer 114
QY 355 ACAGACCCAGGAGAGCAAACTCTCACTATTGAACTGTCTATCATATCATATTCCTCC 414
Db 115 TyrSerGluAspAlaLysValSerThrLeuLysAspIleProHisProSerTyr--- 133
QY 415 ACCAAGAAACCAATGACATGATATGCTTTTGAAGATGCTGGAGCTTCCAAATT 474
Db 134 LeuGlnGluGlySerGlnGlyAspIleAlaLeuLeuGlnLeuSerArgProIleThrPhe 153
QY 475 GGCCACTTTTGGGGCCCATATCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGCTTT 534
Db 154 SerArgTyrIleArgProIleCysLeuProAlaAlaAsnAlaSerPheProAsnGlyLeu 173
QY 535 ATTTGTACAACTCAGGCTGGGGCCCTTAACCTGAAGTGGGCTCCTC-----TCACAA 588
Db 174 HisCysThrValThrGlyTyrGlyHisValAlaProSerValSerLeuLeuThrProLys 193
QY 589 GTCTTCAGGAAGTGAATCTGCTATTTGACCTGGGAAGAGTGTGGCAGCTCTGTTA 648
Db 194 ProLeuGlnGlnLeuGluValProLeuIleSerArgGluThrCysAsnCysLeuTyrAsn 213
QY 649 ACATAAAGAGGCC-----ATCAGTGGGAAGACCTTTCTTTGTCACAGGTTT 696
Db 214 IleAspAlaLysProGluGluProHisPheValGlnGluAspMetValCysAlaGlyTyr 233
QY 697 CTGTATGGAGGAGAGACCATCTCAGGAGATTCAGGAGTTTCACTCATGTGCGCGAAT 756
Db 234 ValGluGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuSerCys---Pro 252
QY 757 AAGAAAGGCGCTGGACTCTGGCTGTGTGACTTCTCTGGGTTGGGCTGTGTCAGGC 816
Db 253 ValGluGlyLeuTyrTyrLeuThrGlyIleValSerTyrPheLysAspAlaCysGlyAla--- 271
QY 817 TGGAGAAACAATGTGAGAAAGTGTCAAGGATCCCTGGGATCTTCACAGACATTAGT 876
Db 272 ---ArgAsnArg-----ProGlyValTyrThrLeuAlaSer 282
QY 877 AAGTGTCTTTCGGATCCACGACACATC 906
Db 283 SerTyrAlaSerTrpIleGlnSerLysVal 292
```

RESULT 9

```
US-09-027-337-3
; Sequence 3, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in
; TITLE OF INVENTION: Breast and Ovarian Carcinomas
```

```
; FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027,337B
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 3
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of hepsin (Heps)
; OTHER INFORMATION: homologous to similar domain in TAGD-15
US-09-027-337-3
```

Alignment Scores:

Pred. No.:	1.03e-39	Length:	256
Score:	437.50	Matches:	93
Percent Similarity:	51.35%	Conservative:	40
Best Local Similarity:	35.91%	Mismatches:	103
Query Match:	26.23%	Indels:	23
DB:	2	Gaps:	7

US-09-735-713A-1 (1-921) x US-09-027-337-3 (1-256)

```
QY 163 CGCATTTCTTGGAGAGCAAGCTGGAGAGGTTCTTATCCCTGGCAGGTATCTCGAA 222
Db 1 ArgIleValGlyGlyArgAspThrSerLeuGlyArgTyrProTyrGlnValSerLeuArg 20
QY 223 CAAGCCAGAGCAATATTTGTGGAGAGCATGCTCTACCCAGTGGGTGATCAGCGG 282
Db 21 TyrAspGlyAlaHisLeuCysGlyGlySerLeuLeuSerGlyAspThrValLeuThrAla 40
QY 283 GGTCACTGATGTCACAAACAGAAAC---ATTGTGTCTACTTTGAATGTTACTGCTGGAGAG 339
Db 41 AlaHisCysPheProGluArgAsnArgValLeuSerArgTyrPheAlaGlyAla 60
QY 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTGCATCAT 399
Db 61 -----ValAlaGlnAlaSerProHisGlyLeuGlnGlyValGlnAlaValTyr 78
QY 400 CATCCACATTTC-----TCCACCAAGAAACCAATGCATGATGATTGTC 444
Db 79 HisGlyGlyTyrLeuProPheArgAspProAsnSerGluGluAsnSerAsnAspIleAla 98
QY 445 CTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCGCCATATCTCTCCA 504
Db 99 LeuValHisLeuSerSerProLeuProLeuThrGluTyrIleGlnProValCysLeuPro 118
QY 505 GAGCTCGGGGACAAATTTGAGGCTGTTTATTTTGTACAACTGCAGGCTGGGCGCGCTTA 564
Db 119 AlaAlaGlyGlnAlaLeuValAspGlyLysIleCysThrValThrGlyTyrGlyAsnThr 138
QY 565 ACTGAAGTGGCGCTCTCTCACAAAGTCTTCCAGGAAGTGAATGCTGCTATTGTTGACCTGG 624
Db 139 GlnTyrTyrGlyGlnGlnAlaGlyValLeuGlnGluAlaArgValProIleLeuSerAsn 158
QY 625 GAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTCTT 684
Db 159 AspValCysAsnGlyAlaAspPhe---TyrGlyAsnGlnIleLysProLysMetPhe--- 176
QY 685 TGCACAGGTTTCTCTGATGAGGAGAGACGCGATGTCAGGGAGATTCAGGAGGTTCACTC 744
Db 177 CysAlaGlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSerGlyGlyProPhe 196
QY 745 ATGTGCGCG-----AATAAGAAAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
Db 197 ValCysGluAspSerIleSerArgThrProArgTyrProArgLeuCysGlyIleValSerTrp 216
QY 796 GGTTTGGGCTGTGCTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCCCT 855
Db 217 GlyThrCysAlaLeuAlaGlnLys-----Pro 226
QY 856 GGGATCTTCACAGACATTAGTAAGATGCTTCTGATCCACGACACATCCAACT 912
Db 856 GGGATCTTCACAGACATTAGTAAGATGCTTCTGATCCACGACACATCCAACT 912
```

10/26/99

Db 227 GlyValTyrThrLysValSerAspPheArgGluTrpIlePheGlnAlaIleLysThr 245
RESULT 10
US-09-644-600-3
; Sequence 3, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hiroto
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIE/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 3
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Hepsin
US-09-644-600-3
Alignment Scores:
Pred. No.: 1.03e-39 Length: 256
Score: 437.50 Matches: 93
Percent Similarity: 51.35% Conservative: 40
Best Local Similarity: 35.91% Mismatches: 103
Query Match: 26.23% Indels: 23
Gaps: 7
US-09-735-713A-1 (1-921) x US-09-644-600-3 (1-256)
QY 163 CGCATTCCTGGAGGAGCAAGTGGAGAGGTCCTCTCCCTGGCAGGTATCTCGAAA 222
Db 1 ArgIleValGlyGlyArgAspThrSerLeuGlyArgTrpProTrpGlnValSerLeuArg 20
QY 223 CAAGGCGAAGCATATTGTGGAGGAGCATGCTCCACACAGTGGGTGATCAGCGCG 282
Db 21 TyrAspGlyAlaHisLeuCysGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAla 40
QY 283 GCTCACTGTCATGCAACACAGAAC--ATTGTGCTACTTTGTATGTACTGTCTGGAGAG 339
Db 41 AlaHisCysPheProGluArgAsnArgValLeuSerArgTrpArgValPheAlaGlyAla 60
QY 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTGATCATA 399
Db 61 -----ValAlaGlnAlaSerProHisGlyLeuGlnLeuGlyValGlnAlaValValTyr 78
QY 400 CATCCACATTC-----TCCACCAAGAAACCAATGGACTATGATATTGCC 444
Db 79 HisGlyGlyTyrLeuProPheArgAspProAsnSerGluGluAsnSerAsnAspIleAla 98
QY 445 CTTTGTGAAGATGGTGGAGCCTTCCATTTGGCCACATTTGTGGGGCCCATATGCTTCCA 504
Db 99 LeuValHisLeuSerSerProLeuProLeuThrGluTyrIleGlnProValCysLeuPro 118
QY 505 GAGTGGGGGAGCAATTTGAGGCTGGTTTATTTGTACAACTCGAGCTGGGGCCGCTTA 564
Db 119 AlaAlaGlyGlnAlaLeuValAspGlyLysIleCysThrValThrGlyTrpGlyAsnThr 138
QY 565 ACTGAAGTGGCTCTCTCACAAGTCTTCGAGGAAGTGAATCGCCTATTTGACCTGG 624
Db 139 GlnTyrTyrGlyGlnGlnAlaGlyValLeuGlnGluAlaArgValProIleIleSerAsn 158
QY 625 GAAGAGTGTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTCTT 684
Db 159 AspValCysAsnGlnAlaAspPhe---TyrGlyAsnGlnIleLysProLysMetPhe--- 176

QY 685 TGCACAGGTTTCTCTGATGGAGGAGAGACGCATGTCCAGGAGATTCAGGAGGTTCACTC 744
Db 177 CysAlaGlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSerGlyGlyProPhe 196
QY 745 ATGTGCGCG-----AATAGAAAGGGCTCGACTCTCGCTGTGTGACTTCTCTGG 795
Db 197 ValCysGluAspSerIleSerArgTrpProArgTrpArgLeuCysGlyIleValSerTrp 216
QY 796 GGTTTGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGGAAAGTGCATCAAGATCCCT 855
Db 217 GlyThrGlyCysAlaLeuAlaGlnLys-----Pro 226
QY 856 GGGATCTTCACAGACATTAGTAAGTGTCTTCTCGATCCACGAAACACATCCAACT 912
Db 227 GlyValTyrThrLysValSerAspPheArgGluTrpIlePheGlnAlaIleLysThr 245
RESULT 11
US-09-386-642-12
; Sequence 12, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-12
Alignment Scores:
Pred. No.: 1.65e-39 Length: 319
Score: 436.00 Matches: 90
Percent Similarity: 51.74% Conservative: 44
Best Local Similarity: 34.75% Mismatches: 107
Query Match: 26.14% Indels: 18
Gaps: 6

US-09-735-713A-1 (1-921) x US-09-386-642-12 (1-319)
QY 148 TTTAAACATTTTCAGTCGCATTTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGG 207
Db 37 PheAspAspAspAspLysIleValGlyTyrAlaLeuGluAlaGlyGlnTrpProTrp 56
QY 208 CAGTATCTCTGAACAAAGGAGACCATATTTGTGGAGAGCATCTGCTCACCACAG 267
Db 57 GlnValSerIleThrTyrGluGlyValHisValCysGlyGlySerLeuValSerGluGln 76
QY 268 TGGTGTATCAGCGGCTCAGTCATTTGCAACAGAAACATTTGTCTACTTTGAATGTT 327
Db 77 TrpValLeuSerAlaAlaHisCysPheProSerGluHisLysGluAlaTyrGluVal 96
QY 328 ACTGCTGGAGATGATGACTTAAAGCCAGACAGACCCAGGAGCAAACTCTCACTATTGAA 387
Db 97 LysLeuGlyAlaHisGlnLeuAspSerTyrSerGluAspAlaLysValSerThrLeuLys 116
QY 388 ACTGTCATCATCATTCACATTTCTCCACCAAGAAACCAATGGACTATGATTTGCCCTT 447
Db 117 AspileProHisProSerTyr---LeuGlnGluGlySerGlnGlyAspIleAlaLeu 135
QY 448 TTGAAGATGCTGGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATGCTTCCAGAG 507
Db 136 LeuGlnLeuSerArgProIleThrPheSerArgTrpIleArgProIleCysLeuProAla 155

QY	106	ACTGTGGGCAGAGTCTGCTTAAGCTACAGCCTTGGAAATATTTTAAACATTTTCAGTCGC	165
		::: :::	:::
Db	148	ThrCysGlyGlu-----LeuTyThrArgSerAsnArg	158
QY	166	ATCTCTGGAGGACCAAGTGGAGRAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACA	225
		:	
Db	159	IleValGlyGlyHisSerThrGlyPheGlySerHisProtrpGluAlaLeuIleuLys	178
QY	226	-----AGCCAGAAGCATATTTGTGGAGGAAGCATGCTCTCACACAGTGGGTG	273
		:::	
Db	179	SerGlyPheLeuSerLysLysLeuSerCysGlyAlaLeuValSerAspArgtrpVal	198
QY	274	ATCAGCGCGCTCAGCTGTCACAAACAAACATTTGCTCTACTTTGAATCTTACTGCT	333
Db	199	IleThrAlaAlaHisCysValIalThr---ThrProAsnSerAsnLeuLysValArgLeu	217
QY	334	GGAGATATGACTTAAGCCAGACAGACCCA-----GGAGACCAACTCTCTACTATTGAA	387
		: ::	
Db	218	GlyGluTrpAspValArgAspHisAspGluArgLeuAsnHisGluLutyrAlaIleGlu	237
QY	388	ACTGTCTCATCATCCACATTTCTCCACCAAGAACAATGGACTAT-----GATATT	441

; INFORMATION FOR SEQ ID NO: 33:

; INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-004-731-33

Alignment Scores:
Pred. No.: 4,4e-39 Length: 400
Score: 432.50 Matches: 101
Percent Similarity: 53.85% Conservative: 46
Best Local Similarity: 37.00% Mismatches: 91
Query Match: 25.93% Indels: 35
DB: 4 Gaps: 11

US-09-735-713A-1 (1-921) x US-09-004-731-33 (1-400)

```
QY 106 AGTTGGGCGAGAGTCTGTTAAGGTACAGCCTTGAATTTTAAACATTTTTCAGTCGC 165
Db 148 ThrCysGlyGlu-----LeuTyrThrArgSerAsnArg 158

QY 166 ATTCTGGAGGACCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAA 225
Db 159 IleValGlyGlyHisSerThrGlyPheGlySerHisProtrpGlnAlaLeuLys 178

QY 226 -----AGGCAGAACATATTGTGGAGGAGCATCTCTCACCACAGTGGGTG 273
Db 179 SerGlyPheLeuSerLysLysLeuSerCysGlyGlyAlaLeuValSerAspArgTrpVal 198

QY 274 ATCAGCGCGCTCAGTCATGTCATGCAACAGAACATTTGTGCTACTTTGAATGTACTGCT 333
Db 199 IleThrAlaAlaHisCysValAlaThr---ThrProAsnSerAsnLeuLysValArgLeu 217

QY 334 GGAGACTATGACTTAAGCCAGACAGACCCCA-----GGAGAGCAAACTCTCATTATGAA 387
Db 218 GlyGluTrpAspValArgAspHisAspGluArgLeuAsnHisGluLufyAlaIleGlu 237

QY 388 ACTGTCATCATCATACATCCACATTTCTCCACCAAGAAACATGACATAT-----GATATT 441
Db 238 ArgGlyGluValHisProSerTyrSer-----ProThrAspPheArgAsnAspVal 254

QY 442 GCCCTTTTGAAGTGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTT 501
Db 255 AlaLeuValLysLeuAspArgThrValIlePheLysGlnHisIleLeuProValCysLeu 274

QY 502 CCAGAGCTGGCGGAGCAATTTGAGGTGGTTTATTGTACAACTCGAGGCTGGGGCCGC 561
Db 275 ProHis---LysGlnMetLysLeuAlaGlyLysMetAlaThrValAlaGlyTrpGlyArg 293

QY 562 TTAAGTGAAGGT---GGCGCCCTCTCACAGTCTTGCAGGAGTGAATCTGCCTATTG 618
Db 294 ThrArgHisGlyGlnSerThrValProAlaValLeuGlnGluValAspValGluValIle 313

QY 619 ACCTGGGAGAGTGTGTGGCAGCTCTGTTAACACTA---AAGAGGCCCATCAGTGGGAAG 675
Db 314 ProAsnGluArgCysGlnArgTppPheArgAlaAlaGlyArgGluThrIleHisAsp 333

QY 676 ACCTTTCTTTCACAGGTTTCTGTGAGGAGGAGACCCATGTCAGGAGATTGAGGA 735
Db 334 ValPheLeuCysAlaGlyTyrLysGluGlyArgAspSerCysGlnGlyAspSerGly 353

QY 736 GGTTCCTACTATGTCGCGGATAAGAAGGGCCCTGGACTCTGCTGGTGTGACTTCTG 795
Db 354 GlyProLeuIleMetGln---IleGluGlyArgArgThrLeuValGlyLeuValSerTrp 372

QY 796 GGTTCGCTGTGCTGAGGCTGGAGAACAAATGTGAGGAAAAAGTATCAAGGATCCCT 855
Db 373 GlyIleGlyCysGlyArg-----GluHisLeuPro 382

QY 856 GGATCTTCACACATAGTAGTAAGTCTTCTCGATC 894
Db 383 GlyValTyrThrAsnIleGlnLysPheIleProTrpIle 395
```

RESULT 15

US-08-749-699-30 : Sequence 30, Application US/08749699
Patent No. 6210920
GENERAL INFORMATION:
APPLICANT: Wu Hunter Shirley
APPLICANT: Stiegler Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-749-699-30

Alignment Scores:
Pred. No.: 4,4e-39 Length: 400
Score: 432.50 Matches: 101
Percent Similarity: 53.85% Conservative: 46
Best Local Similarity: 37.00% Mismatches: 91
Query Match: 25.93% Indels: 35
DB: 4 Gaps: 11

US-09-735-713A-1 (1-921) x US-08-749-699-30 (1-400)

```
QY 106 AGTTGGGCGAGAGTCTGTTAAGGTACAGCCTTGAATTTTAAACATTTTTCAGTCGC 165
Db 148 ThrCysGlyGlu-----LeuTyrThrArgSerAsnArg 158

QY 166 ATTCTGGAGGACCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAA 225
Db 159 IleValGlyGlyHisSerThrGlyPheGlySerHisProtrpGlnAlaLeuLys 178

QY 226 -----AGGCAGAACATATTGTGGAGGAGCATCTCTCACCACAGTGGGTG 273
Db 179 SerGlyPheLeuSerLysLysLeuSerCysGlyGlyAlaLeuValSerAspArgTrpVal 198

QY 274 ATCAGCGCGCTCAGTCATGTCATGCAACAGAACATTTGTGCTACTTTGAATGTACTGCT 333
Db 199 IleThrAlaAlaHisCysValAlaThr---ThrProAsnSerAsnLeuLysValArgLeu 217

QY 334 GGAGACTATGACTTAAGCCAGACAGACCCCA-----GGAGAGCAAACTCTCATTATGAA 387
Db 218 GlyGluTrpAspValArgAspHisAspGluArgLeuAsnHisGluLufyAlaIleGlu 237

QY 388 ACTGTCATCATCATACATCCACATTTCTCCACCAAGAAACATGACATAT-----GATATT 441
```

```
Db 238 ArgLysGluValHisProSerTyrSer-----ProThrAspPheArgAsnAspVal 254
QY 442 GCCCTTTTGAAGAGTGGCTGGAGCCTTCCAAATTGGCCACTTTGTGGGCCCATATGCTT 501
Db 255 AlaleuValLysLeuAspArgThrValIlePheLysGlnHisIleLeuProValCysLeu 274
QY 502 CCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTTGTACAACTGCAGGCTGGGGCGC 561
Db 275 ProHis---LysGlnMetLysLeuAlaGlyLysMetAlaThrValAlaGlyTrpGlyArg 293
QY 562 TTAAGTGAAGT---GGCGTCTCTCACAACTCTTCAGGAGTCAATCTCCCTATTTTG 618
Db 294 ThrArgHisGlyGlnSerThrValProAlaValLeuGlnGluValAspValGluValIle 313
QY 619 ACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTA---AAGAGGCCCATCAGTGGGAAG 675
Db 314 ProAsnGluArgCysGlnArgTrpPheArgAlaAlaGlyArgArgGluThrIleHisAsp 333
QY 676 ACCTTTCTTGCACAGGTTTTCCTGATGGAGGGAGAGACGATGTCAGGGAGATTTCAGGA 735
Db 334 ValPheLeuCysAlaGlyTyrLysGluGlyGlyArgaspSerCysGlnGlyAspSerGly 353
QY 736 GGTTCACATGTCGCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCCTGG 795
Db 354 GlyProLeuIleMetGln---IleGluGlyArgArgThrLeuValGlyLeuValSerTrp 372
QY 796 GGTGTGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGTGCATCAAGGATCCCT 855
Db 373 GlyIleGlyCysGlyArg-----GluHisLeuPro 382
QY 856 GGGATCTTCACAGACATTAGTAAAGTGTTCCTGGATC 894
Db 383 GlyValTyrThrAsnIleGlnLysPheIleProTrpIle 395
```

Search completed: March 25, 2003, 07:01:00
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: March 25, 2003, 06:31:51 ; Search time 42.5 Seconds
(without alignments)
4166.579 Million cell updates/sec

Title: US-09-735-713A-1
Perfect score: 1668
Sequence: 1 atgagtctcaaatgcttat.....acatccaaactggtactaa 921

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-O/cgn2_1/USPTO_SPOOL/US09735713/runat_18032003_124354_23139/app_query.fasta_1.1095
-DB=PIR_73 -OFFMT=fastan -SUFFIX=ipr -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713 @CGN.1.1.26.@runat_18032003_124354_23139 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGUEURY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEROUT=120
-WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	690.5	41.4 1004 2	T30338
2	626	37.5 1524 2	oviductin (EC 3.4. polyprotein - Afri
3	486.5	29.2 786 1	A47547 serine proteinase
4	466	27.9 638 1	KQHUP plasma kallikrein
5	464	27.8 638 1	KQMSPL plasma kallikrein
6	446	26.7 638 1	KQRTPL plasma kallikrein
7	442.5	26.5 1035 1	A43090 enteropeptidase (E
8	441	26.4 343 1	A57014 prostasin (EC 3.4. enteropeptidase (E
9	440.5	26.4 1034 1	A53663 enteropeptidase (E
10	440	26.4 417 1	S00845 hepsin (EC 3.4.21.
11	438	26.3 416 1	S33777 hepsin (EC 3.4.21.
12	433	26.0 593 2	S45281 coagulation factor
13	431.5	25.9 1019 1	A56318 enteropeptidase (E
14	421	25.2 855 2	JC7731 membrane-bound arg

15	421	25.2	855	2	JC7775	membrane type-seri
16	420.5	25.2	1047	2	A55617	masquerade precurs
17	419.5	25.1	237	1	TRCY1	trypsin (EC 3.4.21
18	419	25.1	264	2	I38136	chymotrypsin-like
19	415.5	24.9	275	2	S40007	trypsin (EC 3.4.21
20	415	24.9	625	1	KFH01	coagulation factor
21	414.5	24.9	615	1	KFH012	chymotrypsin (EC 3
22	413.5	24.8	263	2	A21195	chymotrypsin (EC 3
23	409	24.5	263	2	A31299	chymotrypsin (EC 3
24	408.5	24.5	415	1	A34170	acrosin (EC 3.4.21
25	408	24.5	275	2	S40005	trypsin (EC 3.4.21
26	407.5	24.4	274	2	S35339	trypsin (EC 3.4.21
27	406	24.3	254	1	TRW33Y	trypsin-like prote
28	405.5	24.3	603	2	S28941	coagulation factor
29	405	24.3	265	2	T15451	hypothetical prote
30	404	24.2	407	1	KFB07	coagulation factor
31	402.5	24.1	275	2	C35863	trypsin (EC 3.4.2
32	402	24.1	245	1	KYBOB	chymotrypsin (EC 3
33	400	24.0	263	1	KYRTB	chymotrypsin (EC 3
34	400	24.0	275	2	A32410	trypsin (EC 3.4.2
35	399.5	24.0	248	2	S55066	trypsin (EC 3.4.21
36	397.5	23.8	248	2	S55067	trypsin (EC 3.4.21
37	397.5	23.8	461	1	JX0210	protein C (activat
38	397.5	23.8	812	1	PLBO	plasmin (EC 3.4.21
39	396.5	23.8	461	1	KFHU	coagulation factor
40	396	23.7	267	2	S40006	trypsin (EC 3.4.21
41	395	23.7	229	1	TRDFS	trypsin (EC 3.4.21
42	395	23.7	1113	2	JE0315	low-density lipopr
43	394	23.6	263	2	S47537	chymotrypsin (EC 3
44	393.5	23.6	275	2	B35863	trypsin (EC 3.4.2
45	392.5	23.5	269	2	B32410	mastocytoma protei

ALIGNMENTS

RESULT 1
T30338
Oviductin (EC 3.4.21.-) - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 01-Dec-2000
C:Accession: T30338; A40242
R:Lindsay, L.L.; Wieduwilt, M.J.; Hedrick, J.L.
Biol. Reprod. 60, 989-995, 1999
A:Title: Oviductin, the Xenopus laevis oviductal protease that processes egg envelop
composed of two protease and several CUB domains.
A:Reference number: Z20830; MUID:99184825; PMID:10084976
A:Accession: T30338
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1004 <LIN>
A:Cross-references: EMBL:U81291; NID:G1754713; PID:G1754714; PIDN:AAB53972.1
R:Hardy, D.M.; Hedrick, J.L.
Biochemistry 31, 4466-4472, 1992
A:Title: Oviductin purification and properties of the oviductal protease that proces
A:Reference number: A40242; MUID:92256375; PMID:1581303
A:Accession: A40242
A:Molecule type: protein
A:Residues: 46-73 <HAR>
C:Comment: This protease is found in oviductal secretory granules and is secreted to
C:Superfamily: trypsin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase

Alignment Scores:
Pred. No.: 1.86e-55 Length: 1004
Score: 690.50 Matches: 132
Percent Similarity: 60.33% Conservative: 49
Best Local Similarity: 44.00% Mismatches: 110
Query Match: 41.40% Indels: 9
DB: 2 Gaps: 4

US-09-735-713A-1 (1-921) x T30338 (1-1004)

QY 22 ACAGGAACAAGCTGATTTTACTAGGAATAGTCTTTTTCACACAGGTAATCTGCA 81

```
Db      3 ThrArgAsn-----LeuLeuLeuGlySerIleLeuLeuSerLeuAlaValLysGly 19
QY      82 RCCTTTCCGTCCTCCAAAGCTCCAGTGTGGGACAGT---CTGGTTAAGGTACAGCCT 138
Db      20 AspProGlyProHisArgGlyAlaAraGlyValSerProLeuGlySerAlaThrGlu 39
QY      139 TGAATATTATTAACATTTTCACTCCGATCTTGGAGGAGCCAAAGTGGAGAGGCTTCC 198
Db      40 LeuAsnTyr-----LeuSerArgIleValGlyArgGluSerLysGlyGln 56
QY      199 TATCCCTGGCAGGTATCTCTGAACAAAGCAGAAGCATATTTGTGGAGGAAGCATCGTC 258
Db      57 HisProIleThrValSerLeuLysArgAsnGlyLysHisPheCysGlyThrLeuVal 76
QY      259 TCACCACAGTGGGTATCAGCGGGCTCACTGATTCGAAACAGAAACATTTGTGCTACT 318
Db      77 SerHisCysHisValLeuThrAlaAlaHisCysLeuLeuAspArgAsnValLysLeuTyr 96
QY      319 TTGAATGTTACTGCTGGAGAGTATGACTTAAGCCACAGACAGCCAGGAGAGCAAACTCTC 378
Db      97 MetArgValTyrIleGlyGluTyrAspGlnIleLeuLysGluThrGluGlnMetPhe 116
QY      379 ACTATTGAACACTGTCATCATCACATCATTTCTCCACCAAGAAACCAATGGACTATGAT 438
Db      117 ArgValIleGluIlePheLysHisProAsnPheAsnGlnSerGlnProMetAsnTyrAsp 136
QY      439 ATTGCCCTTTTGAAGTGGCTGAGCGCTTCCAAATTTGGCCACTTTTGGGGCCCATATGT 498
Db      137 ValAlaValLeuLeuLeuAspGlySerValThrPheAspGluAsnIleGlnProAlaCys 156
QY      499 CTTCACAGAGCTGGCGGAGCAATTTGAGGCTGGTATTATTTGTAACAACCTGACGCTGGGC 558
Db      157 LeuProAsnProAspAspValPheGluProGlyAspLeuCysValThrLeuGlyTrpGly 176
QY      559 CGTTTAACGAAGTGGCTCTCTCACAGTCTTCAGAGTCTTCAGAGAAAGTAACTGCTATTG 618
Db      177 HisLeuThrGluAsnGlyIleLeuProValValLeuGlnGluValTyrLeuProIleVal 196
QY      619 ACCTGGGAAGAGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACC 678
Db      197 AspLeuSerSerCysLeuHisValMetSerAlaLeuLysGlyThrValValSerSerTyr 216
QY      679 TTTCTTTGCACAGGTTTCTGATGAGGAGGAGACGCATATGTCAGGAGATTCAGGAGGT 738
Db      217 IleValCysAlaGlyPheProGluGlyGlyLysAspAlaCysGlnGlyAspSerGlyGly 236
QY      739 TCACATCATGTGCCGAATAAGAAGGGCTGCACTCTGGCTGGTGTGACTTCTCTGGGT 798
Db      237 ProLeuLeuCysGlnArgArgHisGlySerTrpValLeuHisGlyLeuThrSerTrpGly 256
QY      799 TTGGGCTGTGTCGAGGCTGGAGAAACAATGTG-----AGGAAAGTGATCAAGGATCC 852
Db      257 MetGlyCysGlyArgSerTrpLysAsnAsnValPheLeuProHisAsnArgLysGlySer 276
QY      853 CTGGGATCTTCACAGACATTAGTAAGTCTTTCTTCGATCCACCAAGAACATCCAAACT 912
Db      277 ProGlyIlePheThrAspIleGlnLysLeuLeuGlyTrpValSerSerGlnLeuAsnThr 296

RESULT 2
T30337
polyprotein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30337
R:Yang, J.-C.; Lindsay, L.L.; Hedrick, J.L.
submitted to the EMBL Data Library, March 1998
A:Description: cDNA cloning of ovochymase, a chymotrypsin-like protease released from xe
A:Reference number: Z20829
A:Accession: T30337
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1524 <YAN>
```

```
A:Cross-references: EMBL:U81290; NID:g2981640; PID:g2981641; PIDN:AAC24717.1
Alignment Scores:
Pred. No.:      1.8e-49      Length:      1524
Score:          626.00      Matches:    128
Percent Similarity: 57.89%      Conservative: 37
Best Local Similarity: 44.91%      Mismatches: 110
Query Match:      37.53%      Indels:     10
DB:              2          Gaps:         4

US-09-735-713a-1 (1-921) x T30337 (1-1524)
QY      70 GGTAAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTGTGGGAGAGTCTGGTTAAG 129
Db      32 GlyGluThrAlaGluLeuLysCysGlyThrArgProGluIleGlyAsp----- 47
QY      130 GTACAGCCTTGGAAATTTATTTTAACTATTTTCAGTCGCAATTTTGGAGGAAGCCAAAGTGGAG 189
Db      48 ---GluProAspLeuGluPheThr-----SerArgIleValGlyGlyGlyAspAlaAla 64
QY      190 AAGGTTCCCTATCCCTGGCAGGTATCTCTGAACAAAGCAGCAGCAAGCATATTTGTGGAGGA 249
Db      65 ValGlyGlyGlnProIleThrValSerLeuLysLeuAsnGluArgHisIleCysGlyGly 84
QY      250 AGCATCTCTCACACAGTGGGTGATCAGCGCGCTCACTGCATTT-----GCAAAACAGA 303
Db      85 SerIleValArgLysAspMetValValThrAlaAlaHisCysValTyrProValThrGlu 104
QY      304 ACATTGTGTCACCTTGAATGTTACTGTGTGAGAGATGATGACATTAAGCCAGCAGACACCA 363
Db      105 IleLysValSerHisMetThrValIleValGlyGlyTyrAspGlnGlnValMetAspSer 124
QY      364 GGAGACCAACCTCACTATTGAACTGTCATCATCATCATCATTTCTCCACCAAGAAA 423
Db      125 GlnGluGlnSerIleProValSerHisIleGluProHisProAsnTyrArgGlyAspGly 144
QY      424 CCAATGACATGATATTCCTCTTGAAGATGGCTGGAGCCCTTCAATTTGGCCACTTT 483
Db      145 AsnMetGlyTyrAspIleAlaLeuValPheLeuSerLysProIleIlePheGlySerGln 164
QY      484 GTGGGGCCCATATGCTTCCAGAGCTGCGGAGCAGCATTTTTCAGGCTGGTTTATTGTACA 543
Db      165 ValGlnProIleCysLeuProGlnValGlyGlyLysIleGluAlaGlyThrLeuCysVal 184
QY      544 ACTGAGGCTGGGGCGCTTAAGTGAAGTGGCTCTCTCAACAGCTTTCAGGAGAGTG 603
Db      185 SerSerGlyTrpGlyArgLeuGluGluAsnGlyAspLeuSerProValLeuGlnVal 204
QY      604 AATCTCCCTATTGTCACCTGGGAAGAGTGTGTGGAGAGTCTGTAAACACTAAAGAGGCC 663
Db      205 LysLeuProValValAspAsnGlyThrCysHisAlaValLeuGluProIleGlyHisPro 224
QY      664 ATCAGTGGGAAGACCTTTCTTCCACAGGTTTTCCTGATGGAGGAGAGCAGCATGTCAG 723
Db      225 ValLeuAspAspThrMetLeuCysAlaGlyPheProGluGlyMetAspAlaCysGln 244
QY      724 GGAGATTACAGAGTTCCTCATGTGCGGAATAAGAAAGGGCGCTGAGCTCTGGCTGGT 783
Db      245 GlyAspSerGlyGlyProPheValCysArgArgSerGlyValTrpPheLeuAlaGly 264
QY      784 GTGACTTCTCTGGGTTTGGGCTGTGCTGAGGCTGG---AGAAACAATGTGAGGAAAAGT 840
Db      265 CysValSerTrpGlyLeuGlyCysGlyArgSerTrpGlyAlaLysGlnIleIleArgSer 284
QY      841 GATCAAGGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTTCTCCGATCCACGAA 900
Db      285 GlnSerGlySerProAlaIlePheSerArgValSerSerValLeuAspPheLeuArgPro 304
QY      901 CACATCCAAACTGGT 915
Db      305 ProLysLeuThrGly 309

RESULT 3
```


Alignment Scores:

Pred. No.: 9,88e-35 Length: 638
Score: 466.00 Matches: 91
Percent Similarity: 55.47% Conservative: 51
Best Local Similarity: 35.55% Mismatches: 96
Query Match: 27.94% Indels: 18
DB: 1 Gaps: 5

US-09-735-713A-1 (1-921) x KQHUP (1-638)

```
QY 160 AGTCGATCTTGGAGGAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTG 219
Db 389 ThrArgIleValGlyLeuThrAsnSerSerTrpGlyGluTrpProTrpGlnValSerLeu 408
QY 220 AAACAAGG-----CAGAAGCATATTGTGGAGGAAGCATCTCTCCACACAGTGG 270
Db 409 GlnValLysLeuThrAlaGlnArgHisLeuCysGlyGlySerLeuIleGlyHisGlnTrp 428
QY 271 GTGATCAGCGCGCTCAGTCGATTCGAAACAGAAACATTTGTCTACTTTGAATGTTACT 330
Db 429 ValLeuThrAlaAlaHisCysPheAspGlyLeuProLeuGlnAspValTrpArgIleTyr 448
QY 331 GCTGGAGATGACTTAAGCCAGACAGACCAGGAGCAAACTCTCACTATTGAAACT 390
Db 449 SerGlyIleLeuAsnLeuSerAspIleThrLysAspThrProPheSerGlnIleLysGlu 468
QY 391 GTCATCATATCCACATTTCTCCACCAAGAACCAATGGACTATGATATTGCCCTTTTG 450
Db 469 IleIleIleHisGlnAsnTrpLysValSerGluGly---AsnHisAspIleAlaLeuIle 487
QY 451 AGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGGGGCCCATATGCTTCCAGAGCTG 510
Db 488 LysLeuGlnAlaProLeuAsnTrpThrGluPheGlnLysProIleCysLeuProSerLys 507
QY 511 CGGAGCAATTTGAGGCTGTTTATTCTTACAAGCTGCGAGCTGGGGCCGCTTAAGTGA 570
Db 508 GlyAspThrSerThrIleTrpThrAsnCysTrpValThrGlyTrpGlyPheSerLysGlu 527
QY 571 GGTGGCTCCTCTCACAGCTTTGAGGAAGTGAATCTGCCTATTGACCTGGGAAGAG 630
Db 528 LysGlyGluIleGlnAsnIleLeuGlnLysValAsnIleProLeuValThrAsnGluGlu 547
QY 631 TGTGTGGAGCTCTGTATACACTAAGAGGCCCATCAGTGGGAAGACCTTTCTTGGACA 690
Db 548 CysGlnLysArgTrpGlnAspTrpLys-----IleThrGlnArgMetValCysAla 564
QY 691 GGTTCCTTCATGGAGGAGAGCAGCATCTCAGGAGGATTCAGAGGTTTCATCTATGTC 750
Db 565 GlyTrpLysGluGlyLysAspAlaCysLysGlyAspSerGlyGlyProLeuValCys 584
QY 751 CGGAATAAGAAAGGCGCTTGGACTCTGGCTGTGTCACCTCTCTGGGGTTGGCTGTGT 810
Db 585 LysHis---AsnGlyMetTrpArgLeuValGlyIleThrSerTrpGlyGlyCysAla 603
QY 811 CGAGGCTGGAGAACAAATGTGAGGAAAGTATCAGGATCCCTCGGGATCTTCACAGAC 870
Db 604 Arg-----ArgGluGlnProGlyValTrpThrLys 613
QY 871 ATTAGTAAGTCTTTCCTGGATCCAGCAACACATCCAACTGGAAC 918
Db 614 ValAlaGluTrpMetAspTrpIleLeuGluLysThrGlnSerSerAsp 629
```

RESULT 5

KOMSP

Plasma kallikrein (EC 3.4.21.34) precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999

C:Accession: A36557

R:Seidah, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaubien, G.; Brachpapa, L.; Rochemond

DNA Cell Biol. 9, 737-748, 1990

A:Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and compariso

A:Reference number: A36557; PMID:91090844; PMID:2264928

A:Accession: A36557

A:Molecule type: mRNA

A:Residues: 1-638 <SEI>

A:Cross-references: GB:MS5858; NID:q200358; PIDN:AAA63393.1; PID:q200359

A:Note: part of this sequence, including the amino ends of both the heavy and light c

C:Comment: This protein, synthesized in the liver, circulates as a noncovalent comple

C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a

are linked by one or more disulfide bonds.

C:Superfamily: coagulation factor XI; trypsin homology

C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; in

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-390/Product: plasma kallikrein heavy chain #status experimental <HCH>

F:20-109/Domain: apple repeat <AP1>

F:110-199/Domain: apple repeat <AP2>

F:200-289/Domain: apple repeat <AP3>

F:291-380/Domain: apple repeat <AP4>

F:391-638/Product: plasma kallikrein light chain #status experimental <LCH>

F:391-621/Domain: trypsin homology <TRY>

F:21-104/Domain: tryptophan 194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,

F:127,215,308,396,494/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:434,483,578/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.: 1.51e-34 Length: 638
Score: 464.00 Matches: 95
Percent Similarity: 53.49% Conservative: 43
Best Local Similarity: 36.82% Mismatches: 102
Query Match: 27.82% Indels: 18
DB: 1 Gaps: 6

US-09-735-713A-1 (1-921) x KOMSPL (1-638)

```
QY 154 ATTTTCAGTCGATCTTTGGAGGAAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGlyGluTrpProTrpGlnVal 406
QY 214 TCTCTGAACAAGG-----CAGAAGCATATTGTGGAGGAAGCATCTCTCACCA 264
Db 407 SerLeuGlnValLysLeuValSerGlnThrHisLeuCysGlyGlySerIleIleGlyArg 426
QY 265 CAGTGGGTGATCAGCGGCTCAGTCGATTCGAAACAGAAACATCTGCTACTTTGAAT 324
Db 427 GlnTrpValLeuAlaHisCysPheAspGlyIleProTrpProAspValTrpArg 446
QY 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACTATT 384
Db 447 IleTyrGlyGlyIleLeuSerLeuSerGluIleThrLysGluThrProSerArgIle 466
QY 385 GAACTGTCATCATCATCCACATTTCTCCACCAAGAACCAATGGACTATGATATTGCC 444
Db 467 LysGluLeuIleIleHisGlnGluTrpLysValSerGluGly---AsnTrpAspIleAla 485
QY 445 CTTTTCAGAGTGGCTGGAGCCCTCCCAATTTGGCCACTTTGTGGGCCCCATATGCTTCCA 504
Db 486 LeuIleLysLeuGlnThrProLeuAsnTrpThrGluPheGlnLysProIleCysLeuPro 505
QY 505 GAGCTGGGAGCAATTTGAGGCTGTTTATTGTAACACTGACAGCTGGGGCCGCTTA 564
Db 506 SerLysAlaAspThrAsnThrIleThrAsnCysTrpValThrGlyTrpGlyThr 525
QY 565 ACTGAAGTGGCTCCTCTCACAGTCTTTCAGGAGAGTGAATCTGCTATTTTGACCTGG 624
Db 526 LysGluGlnGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
QY 625 GAAGAGTGTGGCAGCTCTGTTAACTAAAGAGAGCCCATCAGTGGGAGACCTTTCTT 584
Db 546 GluGluCysGlnLys-----LysTrpArgAspTrpValIleAsnLysGlnMetIle 562
QY 685 TGCACAGGTTTTCCTGATGGAGGAGAGACGATCTGACGAGGAGATTTCAGAGGTTTCACATC 744
Db 563 CysAlaGlyTrpLysGluGlyGlyThrAspAlaCysLysGlyAspSerGlyGlyProLeu 582
QY 745 ATGTCCCGGAATPAGAAAGGCGCTTGGACTCTGGCTGTGTGACTTCTCTGGGTTGGGC 804
```


859 LeuHisMetAlaSerAsnLeuThrSerProGlnIleGluThrArgLeuIleAspGlnIle 878
394 ATCATACATCCACATTCTCCACCAAGAACCAATGAGCTATGATATTCGCCCTTTGAAG 453
879 ValIleasnProHisTyrAsn---LysArgArgLysAsnAsnAspIleAlaMetMetHis 897
454 ATGGCTCGAGCCCTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGCAGCTGGG 513
898 LeuGluMetLysValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGluAsn 917
514 GAGCAATTTGAGGCTGGTTTTATTGTACAACGTGACAGCTGGGGCCGCTTAACCTGAAGGT 573
918 GlnValPheProGlyArgIleCysSerIleAlaGlyTyrpGlyAlaLeuIleTyrGln 937
574 GGGCTCTCTCACAAAGCTTGCAGGAGAGTGAATCTGCCATTATTGTACCTGGGAAAGAGTGT 633
938 GlySerThrAlaAspValLeuGlnGluAlaAspValProLeuLeuSerAsnGluLysCys 957
634 GTGGCAGCTCTG---TTAAACACTAAAGAGGCCCATCAGTGGGAAAGACCTTT 681
958 GlnGlnGlnMetProGluTyrAsnIleThr-----GluAsnMet 970
682 CTTTGCACAGGTTTCTGATGAGGAGGAGAGCAGCATGTACAGGAGATTTCAGGAGGTTTCA 741
971 ValCysAlaGlyTyrGluAlaGlyGlyValAspSerCysGlnGlyAspSerGlyGlyPro 990
742 CTCATGTGCGCGAATAGAAGGGCGCTGGACTCTGGCTGGTGTGACTCTCTCTGGGGTTTG 801
991 LeuMetCysGln---GluAsnAsnArgTyrLeuLeuAlaGlyValThrSerPheGlyTyr 1009
802 GGCTGTGGTGGAGCTGGAGAAACAATGTGAGGAAAAGTGATCAAGGATCCCTCGGATC 861
1010 GlnCysAlaLeuProAsnArg-----ProGlyVal 1019
862 TTCACAGACATTAGTAAAGTGCTTTCTCGATCCACGACAAACATC 906
1020 TyrAlaArgValProArgPheThrGluTyrPheGlnSerPheLeu 1034
RESULT 8
A57014
proctasin (EC 3.4.21.-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 18-Jun-1999
C:Accession: A57014; A54866
R:Yu, J. X.; Chao, L.; Chao, J.
J. Biol. Chem. 270, 13483-13489, 1995
A:Title: Molecular cloning, tissue-specific expression, and cellular localization of the human proctasin gene
A:Reference number: A57014; MUID:95286644; PMID:7768952
A:Accession: A57014
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-343 <RES>
A:Cross-references: GB:I41351; NID:g862304; PIDN:AAC41759.1; PID:g862305
A:Experimental source: prostate
A:Note: Parts of this sequence were determined by protein sequencing
R:Yu, J. X.; Chao, L.; Chao, J.
J. Biol. Chem. 269, 18943-18948, 1994
A:Title: Proctasin is a novel human serine proteinase from seminal fluid. P
A:Reference number: A54866; MUID:94308140; PMID:8034638
A:Accession: A54866
A:Molecule type: protein
A:Residues: 45-64 <YUA>
C:Genetics:
A:Gene: GDB:PRS58
A:Cross-references: GDB:676446; OMIM:600823
A:Map position: 16p11.2-16p11.2
C:Superfamily: proctasin; hydrolase; trypsin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane prote
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-44,45-343/product: proctasin #status predicted <MAT>
F:33-44/Domain: proctasin light chain #status predicted <CHL>
F:45-343/Domain: proctasin heavy chain #status predicted <CHH>
F:45-281/Domain: trypsin homology <TRY>

F:323-341/Domain: transmembrane #status predicted <TM>1>
F:37-154,70-86,168-244,201-223,234-262/Disulfide bonds: #status predicted
F:85,134,238/Active site: His, Asp, Ser #status predicted
F:159/Binding site: carbohydrate (Asn) (covalent) #status experimental

```
Alignment Scores:
Pred. No.: 1.86e-32 Length: 343
Score: 441.00 Matches: 100
Percent Similarity: 49.16% Conservative: 46
Best Local Similarity: 33.67% Mismatches: 117
Query Match: 26.44% Indels: 34
DB: 1 Gaps: 10

US-09-735-713A-1 (1-921) x A57014 (1-343)
QY 34 CTGATTTTACTAGGATAGTCTTTTGAACRAGTAATTCGCACTCTTCGCTC 93
Db 17 IleLeuLeuTyLeuGlyLeuLeuArgSerGlyThrGlyAlaGluGlyAla----- 33
QY 94 CCAAAAGCTCCAGTGTGTGGCAGAGTCTGTTAAGGTACAGCTTGGAAATTAATTTAAC 153
Db 34 ---GluAlaPro---CysGly-----ValAlaProGln----- 42
QY 154 ATTTTCAGTCGATCTTGGAGGAGCCAAAGTGGAGAGGTCTCTATCCCTGGCAGGTA 213
Db 43 -----AlaArgIleThrGlyGlySerSerAlaValAlaGlyGlnTrpProTrpGlnVal 60
QY 214 TCTTGAACAAAGCCAGAACATATTTGTGGAGGAACATCGTCTCACCAGTGGGTG 273
Db 61 SerIleThrTyGluGlyValHisValCysGlyGlySerLeuValSerGluGlnTrpVal 80
QY 274 ATCAGCGGCTGCTACTGCTTGAACAAAGAACATTTGTGTACTTTGAATGTTACTGCT 333
Db 81 LeuSerAlaAlaHisCysPheProSerGluHisHisLysGluAlaTyGluValLysLeu 100
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGAGCAACTCTCATATTGAACACTGTC 393
Db 101 GlyAlaHisGlnLeuAspSerTySerGluAspAlaLysValSerThrLeuLysAspIle 120
QY 394 ATCATACATCACATTTCTCCACCAAGAAACCAATGGACTATGATATTCCTTTTGAAG 453
Db 121 IleProHisProSerTy---LeuGlnGlySerGlnGlyAspIleAlaLeuLeuGln 139
QY 454 ATGGCTGGAGCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGAGCTGCGG 513
Db 140 LeuSerArgProIleThrPheSerArgTyIleArgProIleCysLeuProAlaAlaAsn 159
QY 514 GAGCAATTTGAGCTGTTTATTGTACAACTGCAGCTGGCGCGCTTAACTGAAGGT 573
Db 160 AlaSerPheProAsnGlyLeuHisCysThrValThrGlyTrpGlyHisValAlaProSer 179
QY 574 GCGGTCCPC-----TCAAGAGTCTTGGAGGAGTAATCTGCTATTTTTCACCTGGAA 627
Db 180 ValSerLeuLeuThrProLysProLeuGlnGlnLeuGluValProLeuIleSerArgGlu 199
QY 628 GAGTGTGGCAGCTCTCTTAACACTAAAGAGGCC-----ATCAGTGGGAG 675
Db 200 ThrCysAsnCysLeuTyAsnIleAspAlaLysProGluGluProHisPheValGlnGlu 219
QY 676 ACCTTTCTTTTCACAGGTTTTCCTGATGGAGGAGACGATGTCAGGAGATTCAGGA 735
Db 220 AspMetValCysAlaGlyTyValGluGlyGlyLysAspAlaCysGlnGlyAspSerGly 239
QY 736 GGTTCACATGTCGCCGAATAAGAAAGCGCTGGAATCTGGTGTGTCATCTCCG 795
Db 240 GlyProLeuSerCys---ProValGluGlyLeuTrpTyThrGlyIleValSerTrp 258
QY 796 GGTGTGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGGAAAGTATCAAGGATCCCT 855
Db 259 GlyAspAlaCysGlyAla-----ArgAsnArg-----Pro 268
QY 856 GGGATCTTCACAGACATTAGTAAAGTCTTCTCGGATCCAGCAACATC 906
Db 856 GGGATCTTCACAGACATTAGTAAAGTCTTCTCGGATCCAGCAACATC 906
```

Db 269 GlyValTyThrLeuAlaSerSerTyThrAlaSerTrpIleGlnSerLysVal 285
RESULT 9
A53663
enteropeptidase (EC 3.4.21.9) precursor - pig
N:Alternate names: enterokinase
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Oct-1994 #sequence_revision 09-Aug-1996 #text_change 18-Jun-1999
C:Accession: A53663
R:Matsushima, M.; Ichinose, M.; Yahagi, N.; Kakel, N.; Tsukada, S.; Miki, K.; Kuroka
J. Biol. Chem. 269, 19976-19982, 1994
A:Title: Structural characterization of porcine enteropeptidase.
A:Reference number: A53663; MUID:94327548; PMID:8051081
A:Accession: A53663
A:Molecule type: mRNA
A:Residues: 1-1034 <MAT>
A:Cross-references: GB:D30799; NID:g505122; PIDN:BAA06459.1; PID:g505123
A:Note: parts of this sequence, including the amino ends of three chains isolated fr
C:Comment: The mechanism of association with the membrane of the intestinal brush bo
ated below) or with amino-terminal myristoylation of the heavy chain.
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and l
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms in
C:Function:
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
A:Pathway: intestinal digestive hydrolase cascade
C:Superfamily: enteropeptidase; Clr/Clis repeat homology; LDL receptor ligand-binding
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymog
F:22-38/Domain: transmembrane #status predicted <TM>
F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>
F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:357-519/Domain: MAM homology <MAM>
F:541-646/Domain: Clr/Clis repeat homology <CLR>
F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical
F:800-1034/Product: enteropeptidase light chain #status predicted <LCH>
F:800-1029/Domain: trypsin homology <TRY>
F:116,147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,90
F:787-911,825-841,925-992,956-971,982-1010/Disulfide bonds: #status predicted
F:840,891,986/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 2.4e-32 Length: 1034
Score: 440.50 Matches: 101
Percent Similarity: 48.95% Conservative: 62
Best Local Similarity: 30.33% Mismatches: 93
Query Match: 26.41% Indels: 77
DB: 1 Gaps: 11

US-09-735-713A-1 (1-921) x A53663 (1-1034)
QY 43 CTACTAGGAATA-----GTCTTTTGAACRAGGTAATCTGCA 81
Db 733 LeuLeuGlyLeuGlyThrGlyAsnSerSerMetProPhePheSerSerGlyGlyGlyPro 752
QY 82 RCTCTTTTCGCTCCCAAGCTCCC----- 105
Db 753 PheValLysLeuAsnThrAlaProAsnGlySerLeuIleLeuThrAlaSerGluGlnCys 772
QY 106 -----AGGTACAGCCTTGGAAATTTTAAACATTTTCAGTCGCGATCTGGTT 126
Db 773 PheGluAspSerLeuIleLeuLeuGlnCysAsnHisLysSerCysGlyLysGlnVal 792
QY 127 -----AGGTACAGCCTTGGAAATTTTAAACATTTTCAGTCGCGAGGAGC 180
Db 793 AlaGlnGluValSerPro-----LysIleValGlyGlyVal 804
QY 181 CAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAAGCAGAGCATATT 240
Db 805 AspSerArgGluGlyAlaTrpProTrpValValAlaLeuTyThrAsnGlyGlnLeuLeu 824
QY 241 TGTGAGGAGCATCGTCTCACCAGGTGGGTGATCAGCGGGCTCAGTCGATTCGATCAAC 300

A:Reference number: S33777; MUID:93305733; PMID:8318546

A:Accession: S33777

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <FAR>

A:Cross-references: EMBL:X70900; NID:g57928; PIDN:CAA50256.1; PID:g57929

C:Superfamily: hepsin; trypsin homology

C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein

F:22-44/Domain: transmembrane #status predicted <TMN>

F:162-399/Domain: trypsin homology <TRY>

F:187-203,290-358,321-337,348-380/Disulfide bonds: #status predicted

F:202,256,352/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	3,62e-32	Length:	416
Score:	438.00	Matches:	99
Percent Similarity:	49.47%	Conservative:	41
Best Local Similarity:	34.98%	Mismatches:	99
Query Match:	26.26%	Indels:	44
DB:	1	Gaps:	10

US-09-735-713A-1 (1-921) x S33777 (1-416)

```
QY 109 TGTGGGAGAGCTGCTTAAAGTACAGCCTTGGAAATTATTTTAAACATTTTCAGTCGCATT 168
|||||
Db 152 CysGlyArgArgLysLeuProVal-----AspArgile 162

QY 169 CTTGGAGAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228
|||||
Db 163 valGlyGlyGlnaspSerSerLeuGlyArgTrpProtrGlnValSerLeuArgTyAsp 182

QY 229 CAGAAGCATATTTCTGGAGGAAGCATGCTCTACACAGAGTGGGTGATCAGCGCGCTCAC 288
|||||
Db 183 GlyThrHisLeuCysGlySerLeuSerGlyAspTrpValLeuThrAlaAlaHis 202

QY 289 TGCATTGCAACACAAAC---ATTGCTGCTACTTTGAATGTTACTGCTGGAGATGAC 345
|||||
Db 203 CysPheProGluArgAsnArgValLeuSerArgTrpArgValPheAlaGlyAla----- 220

QY 346 TTAAGCCAGACAGACCAGGAGAGCAACTCTCACTATTGAACTGTCATCATCATACCA 405
|||||
Db 221 ValAlaArgThrSerProHisAlaValGlnLeuGlyValGlnAlaValIleTyHisGly 240

QY 406 CATTTC-----TCCACCAAGAAACCAATGACTATGATATTGCTCTTTG 450
|||||
Db 241 GlyTyrLeuProPheArgAspProThrIleaspGluAsnSerAsnAspIleAlaLeuVal 260

QY 451 AAGATGGCTGAGCCTTCCCAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTG 510
|||||
Db 261 HisLeuSerSerLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 280

QY 511 CGGGAGCAATTTGAGGCTGTTTATTTGTACAACTGCAGCTGGGGCCGCTTAACGTAA 570
|||||
Db 281 GlyGlnAlaLeuValaspGlyLysValCysThrValThrGlyTrpGlyAsnThrGlnPhe 300

QY 571 GGTGGCTCTCTCACAGTCTTCAGGAAGTGAATCTGCCTATTATTTGACCTGGGAAGAG 630
|||||
Db 301 TyrGlyGlnGlnAlaValValLeuGlnGlnAlaArgValProIleIleSerAsnGluVal 320

QY 631 TGTGTGGCAGCTGTTAAACACTAAAGAGGCC-----ATCAGTGGG 672
|||||
Db 321 Cys-----AsnSerProAspPheTyrGlyAsnGlnIleLysPro 333

QY 673 AAGACCTTCTTCACAGGTTTCTTGATGGAGGAGACGCATGTCAGGGAGATTCA 732
|||||
Db 334 LysMetPhe---CysAlaGlyTyrProGluGlyTyrIleaspAlaCysGlnGlyAspSer 352

QY 733 GGAGGTTCACTCATGTGCCGAATAAG-----AAAGGGCCTGCACCTCGCTGGT 783
|||||
Db 353 GlyGlyHisPheValCysGluAspArgIleSerGlyThrSerArgTrpArgLeuCysGly 372

QY 784 GTGACTTCTCGGGTGTGGCTGGTGGTCGAGGCTGGAGAAACAATGTGAGAAAGTAT 843
|||||
```

```
Db 373 IleValSerTrpGlyThrGlyCysAlaLeu-----AlaArgLys----- 385
QY 844 CAAGGATCCCTGGATCTTCACAGACATTAGTAAAGTGTCTTCTGTGATCCAGACAC 903
|||||
Db 386 -----ProGlyValTyrThrLysValIleaspPheArgGluTrpIlePheGlnAla 402
QY 904 ATCCAAACT 912
|||||
Db 403 IleLysThr 405

RESULT 12
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattive)
C>Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII):
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SHI>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 a
is, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:861210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16,'X',18-19;525-550 <FUJ>
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homol
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plas
F:37-78/Domain: fibronectin type II repeat homology <IP2>
F:88-120/Domain: EGF homology <EGF>
F:125-160/Domain: fibronectin type I repeat homology <FB1>
F:207-287/Domain: kringlike homology <KRK>
F:350-387/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted
```

Alignment Scores:			
Pred. No.:	1,1e-31	Length:	593
Score:	433.00	Matches:	101
Percent Similarity:	52.14%	Conservative:	45
Best Local Similarity:	36.07%	Mismatches:	94
Query Match:	25.96%	Indels:	40
DB:	2	Gaps:	12

US-09-735-713A-1 (1-921) x S45281 (1-593)

```
QY 109 TGTGGCAGAGCTGTTTAAAGTACAGCCTTGGAAATTATTTTAAACATTTTCAGTCGCATT 168
|||||
Db 336 CysGlyGlnArgLeuArgLys-----TrpLeuSerSerLeuAsnArgVal 350

QY 169 CTTGGAGAACCCCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228
|||||
Db 351 ValGlyGlyLeuValAlaLeuProGlyAlaHisProTyrIleAlaLeuTyrTrpAsp 370

QY 229 CAGAAGCATATTTCTGGAGGAAGCATGCTCTACACAGATGGGTGATCAGCGCGCTCAC 288
|||||
Db 371 Gln---HisPheCysAlaGlySerLeuIleAlaProCysTrpValLeuThrAlaAlaHis 389

QY 289 TGCATTGCAACACAAACATTTGCTCTACTTTGAATGTTTACTGCTGGAGAGATGACTTCA 348
|||||
Db 390 CysLeuGlnAsnArgProAlaProLysGluLeuThrValValLeuGlyGlnAspArgHis 409

QY 349 AGCCAGACAGACCAGGAGAGCAAACTCTCACTATTGAAACTGTATCATCATCATCAT 408
|||||
Db 410 AsnGlnSerCysGluGlnCysGlnThrLeuAlaValArgAspTyrArgLeuHisGluAla 429
```


QY 808 GGTGAGGCTGGAGAAACAATGTGTGAGAAAGTGTATCAAGGATCCCTGGGATCTTCTACA 867
Db 996 AlaLeuProAsnArg-----ProGlyValTyrAla 1005
QY 868 GACATTAGTAAGTCTTCTTGGATCCACGACACATC 906
Db 1006 ArgValSerArgPheThrGluTrpIleGlnSerPheLeu 1018
RESULT 14
JC77731
membrane-bound arginine-specific serine proteinase precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 14-Dec-2001
R:Accession: JC77731
R:Kishi, K.; Yamazaki, K.; Yasuda, I.; Yahagi, N.; Ichinose, M.; Tsuchiya, Y.; Athauda, J.; Biochem. 130, 425-430, 2001
A:Title: Characterization of a membrane-bound arginine-specific serine protease from rat
A:Reference number: JC77731; MUID:21421307; PMID:11530019
A:Accession: JC77731
A:Molecule type: mRNA
A:Residues: 1-855 <KIS>
A:Cross-references: DDBJ:AB049189
A:Experimental source: strain Male, 7-week-old
C:Comment: This enzyme, localized mainly on brushborder membranes of the intestine, part
C:Keywords: protein digestion
Alignment Scores:
Pred. No.: 1.48e-30 Length: 855
Score: 421.00 Matches: 96
Percent Similarity: 52.67% Conservative: 42
Best Local Similarity: 36.64% Mismatches: 94
Query Match: 25.24% Indels: 30
DB: 2 Gaps: 8
US-09-735-713A-1 (1-921) x JC7731 (1-855)
QY 160 AGTCGATCTTGGAGAGCAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTG 219
Db 613 AlaArgValValGlyThrAsnAlaAspGluGlyGluTrpProTrpGlnValSerLeu 632
QY 220 AAACAA---AGCAGAACATATTTGTGGAGAGCATCGTCTACACAGTGGGTGATC 276
Db 633 HisAlaLeuGlyGlnGlyHisLeuGlyHisLeuSerProAspTrpLeuVal 652
QY 277 ACGCGGCTACTGCTATTCACAGAACATTTGTCTACTTTGAATCTTACT----- 330
Db 653 SerAlaAlaHisCysPheGlnAspGluThrIlePheLysTyrSerAspHisThrMetTrp 672
QY 331 -----GCTGGAGAGTATGACTTAAGC---CAGACAGACCCAGAGAGCAAACTCTC 378
Db 673 ThrAlaPheLeuGlyLeuLeuAspGlnSerLysArgSerAlaSerGlyValGlnGluHis 692
QY 379 ACTATTGAACGTGTCATCATCATCCATTTCTCCACAGAAACCAATGACTATGAT 438
Db 693 LysLeuLysArgIleThrHisProSerPheAsn---AspPheThrPheAspTyrAsp 711
QY 439 ATTGCCCTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGCCCCATATGT 498
Db 712 IleAlaLeuLeuGluGluLysProAlaGluTyrSerThrValValArgProIleCys 731
QY 499 CTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTTGTAACATGCGAGCTGGGCG 558
Db 732 LeuProAspAsnThrHisValPheProAlaGlyLysAlaIleTrpValThrGlyTrpGly 751
QY 559 CGCTTAACGTGAAGTGGGCTCTCTCACAAGTCTTTCAGAGAGTGAATCTGCCTATTTTG 618
Db 752 HisThrLysGluGlyGlyThrGlyAlaLeuLeuGlnLysGlyGluIleArgValIle 771
QY 619 -----ACCTGGGAAGAGTGTGGCAGCTCTGTTTAACACTAAAGAGGCCCATCAGT 669
Db 772 AsnGlnThrThrCysGluGluLeuLeuProGlnGlnIleThrProArg----- 787
QY 670 GGAAGAGACCTTTCTTGTGACAGAGTTTCTCTGATGGAGGGAGAGAGCGCATGTCTCAGGAGAT 729

Db 788 -----MetMetCysValGlyPheLeuSerGlyGlyValAspSerCysGlnGlyAsp 804
QY 730 TCAGGAGGTTTACATCATGTGCGGGAATAAGAAAGGGCGCTGACTCTCGCTGCTGACT 789
Db 805 SerGlyGlyProLeuSerSerValGluLysAspGlyArgIlePheGlnAlaGlyValVal 824
QY 790 TCCTGGGCTTTGGGCTGTGGTGGAGCTGGAGAAACAATGTGAGAAAGTCAATCAAGGA 849
Db 825 SerTrpGlyGluGlyCys-----AlaGlnArgAsn 834
QY 850 TCCCTGGGATCTTCACAGACATTTAGTAAAGTGTCTTCTGGATCCACGAAACATCCAA 909
Db 835 LysProGlyValTyrThrArgIleProGluValArgAspTrpIleLysGlu-----Gln 852
QY 910 ACTGGT 915
Db 853 ThrGly 854
RESULT 15
JC7775
membrane type-serine protease 1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: JC7775
R:Satomi, S.; Yamasaki, Y.; Tsuzuki, S.; Hitomi, Y.; Iwanaga, T.; Fushiki, T.
Biochem. Biophys. Res. Commun. 287, 995-1002, 2001
A:Title: A role for membrane-type serine protease (MT-SPL) in intestinal epithelial
A:Reference number: JC7775; PMID:11573963
A:Contents: Small intestine
A:Accession: JC7775
A:Molecule type: mRNA
A:Residues: 1-855 <SAT>
A:Cross-references: DDBJ:AB037898
C:Comment: This enzyme, an epithelial-derived, type II integral membrane serine prot
ciated with epithelial migration and/or cell loss not only as an upstream activator
proteins.
A:Gene: mt-spl
A:Map position: basolateral cell surface
Alignment Scores:
Pred. No.: 1.48e-30 Length: 855
Score: 421.00 Matches: 96
Percent Similarity: 52.67% Conservative: 42
Best Local Similarity: 36.64% Mismatches: 94
Query Match: 25.24% Indels: 30
DB: 2 Gaps: 8
US-09-735-713A-1 (1-921) x JC7775 (1-855)
QY 160 ACTGCCATTTGGAGAGCAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTG 219
Db 613 AlaArgValValGlyThrAsnAlaAspGluGlyGluTrpProTrpGlnValSerLeu 632
QY 220 AAACAA---AGCAGAACATATTTGTGGAGAGCATCGTCTACACAGTGGGTGATC 276
Db 633 HisAlaLeuGlyGlnGlyHisLeuGlyHisLeuSerProAspTrpLeuVal 652
QY 277 ACGCGGCTACTGCTATTCACAGAACATTTGTCTACTTTGAATCTTACT----- 330
Db 653 SerAlaAlaHisCysPheGlnAspGluThrIlePheLysTyrSerAspHisThrMetTrp 672
QY 331 -----GCTGGAGAGTATGACTTAAGC---CAGACAGACCCAGAGAGCAAACTCTC 378
Db 673 ThrAlaPheLeuGlyLeuLeuAspGlnSerLysArgSerAlaSerGlyValGlnGluHis 692
QY 379 ACTATTGAACGTGTCATCATCATCCATTTCTCCACAGAAACCAATGACTATGAT 438
Db 693 LysLeuLysArgIleThrHisProSerPheAsn---AspPheThrPheAspTyrAsp 711
QY 439 ATTGCCCTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGCCCCATATGT 498
Db 712 IleAlaLeuLeuGluGluLysProAlaGluTyrSerThrValValArgProIleCys 731
QY 499 CTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTTGTAACATGCGAGCTGGGCG 558
Db 732 LeuProAspAsnThrHisValPheProAlaGlyLysAlaIleTrpValThrGlyTrpGly 751
QY 559 CGCTTAACGTGAAGTGGGCTCTCTCACAAGTCTTTCAGAGAGTGAATCTGCCTATTTTG 618
Db 752 HisThrLysGluGlyGlyThrGlyAlaLeuLeuGlnLysGlyGluIleArgValIle 771
QY 619 -----ACCTGGGAAGAGTGTGGCAGCTCTGTTTAACACTAAAGAGGCCCATCAGT 669
Db 772 AsnGlnThrThrCysGluGluLeuLeuProGlnGlnIleThrProArg----- 787
QY 670 GGAAGAGACCTTTCTTGTGACAGAGTTTCTCTGATGGAGGGAGAGAGCGCATGTCTCAGGAGAT 729

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: March 25, 2003, 05:59:35 ; Search time 182 seconds
(without alignments)
1348.613 Million cell updates/sec

Title: US-09-735-713A-1
Perfect score: 1668
Sequence: 1 atgagctctcaaatgtcttat.....acatccaaactggtactaa 921

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cn2_1/USPTO.spool/US09735713/runat_18032003_124353_23109/app_query.fasta_1.1095
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09735713 @CN1.1.53 @runat_18032003_124353_23109 -NCPU=6 -ICPU=3
-NO_XLPPX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1626	97.5	306	22	AAE04733	Human protease hom
2	1626	97.5	306	23	AAU04760	Human protease PRT
3	1608	96.4	302	22	AAE04734	Human protease hom
4	1356	81.3	556	23	AAU02748	Amino acid sequenc
5	944	56.6	182	22	AAU24275	Toad EST encoded p
6	898	53.8	164	22	AAE04735	Human protease hom
7	567.5	34.0	913	22	ABG24246	Novel human diagno
8	567.5	34.0	1576	22	ABG07870	Novel human diagno
9	567.5	34.0	1576	22	ABG10218	Novel human diagno
10	567.5	34.0	1576	22	ABG14588	Novel human diagno
11	567.5	34.0	1576	22	ABG19887	Novel human diagno
12	567.5	33.7	1031	23	ABF60993	Novel human protei
13	489.5	29.3	802	20	AAU41710	Human PRO618 prote
14	489.5	29.3	802	21	AAU44266	Human PRO618 (UNQ3
15	489.5	29.3	802	21	AAU24052	Human PRO618 prote
16	489.5	29.3	802	23	AAU82755	Amino acid sequenc
17	486	29.1	787	22	ABU71302	Drosophila melanog
18	485.5	29.1	658	22	AAE06934	Human membrane-typ
19	485.5	29.1	802	22	AAE06933	Human membrane-typ
20	476.5	28.6	235	22	AAE06932	Human membrane-typ
21	467	28.0	452	20	AAU41694	Human PRO382 prote
22	467	28.0	452	22	AAU29055	Human PRO polypept
23	467	28.0	453	22	AAE06935	Human membrane-typ
24	467	28.0	453	23	AAE23020	Human trypsin fami
25	467	28.0	454	21	AAU32246	Tumour associated
26	467	28.0	454	22	AAU68911	Human TAGD-12, pro
27	466	27.9	453	21	AAU44250	Human PRO382 (UNQ3
28	456.5	27.4	327	21	AAU72093	Human serine prote
29	456.5	27.4	327	23	AAE17921	Human gene 3 encod
30	456.5	27.4	394	23	ABP41994	Human ovarian anti
31	456.5	27.4	454	23	AAE23024	Human trypsin fami
32	456.5	27.4	454	23	AAU82745	Amino acid sequenc
33	454.5	27.2	248	21	AAU43572	Human cancer assoc
34	450.5	27.0	1128	23	AAU98890	Human protease PRT
35	450.5	27.0	1128	23	AAU82739	Amino acid sequenc
36	446	26.7	243	23	AAU80516	Epithelin-like ser
37	446	26.7	309	23	AAU80531	Epithelin-like ser
38	446	26.7	818	23	AAU82753	Amino acid sequenc
39	444	26.6	414	21	AAU08912	Human secreted pro
40	444	26.6	480	21	AAU08950	Human secreted pro
41	442.5	26.5	798	15	AAU57283	Bovine enterokinas
42	441	26.4	249	23	AAE21441	Human trypsin doma
43	441	26.4	343	23	AAU78547	Human prostatin pr
44	441	26.4	343	23	ABU07285	Amino acid sequenc
45	440	26.4	416	20	AAU43325	Mouse hepsin prote

ALIGNMENTS

RESULT 1
AAE04733
ID AAE04733 standard; Protein; 306 AA.
XX
AC AAE04733;
XX
XX 10-SEP-2001 (first entry)
XX
DE Human protease homologue #1.
XX
KW Human; protease homologue; novel human protein; NHP; therapy;
KW pharmacogenomic application; physiological disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 23
FT /note= "Encoded by CRA"
FT Misc-difference 28

FT /note= "Encoded by RCT"
 XX WO200146407-A1.
 PN 28-JUN-2001.
 PD
 XX
 XX 12-DEC-2000; 2000WO-US33738.
 PF 23-DEC-1999; 99US-0171566.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AF;
 PI WPI; 2001-408641/43.
 XX N-PSDB; AAD09328.
 DR
 XX Polynucleotide encoding novel human protease homologs, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications -
 XX
 XX Claim 2; Page 28-29; 32pp; English.
 PS
 XX The present sequence is novel human protein (NHP),
 CC known as human protease homologue. NHP shares structural similarity
 CC with animal proteases, particularly trypsin-like protease such
 CC as Oviductin, plasminogen activator and human plasma kallikrein
 CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
 CC pharmacogenomic applications. NHP sequences are useful for identifying
 CC agonists, antagonists and modulators and also for producing antibodies
 CC useful in diagnosis, drug screening, clinical trial monitoring and in
 CC treatment of physiological disorders.
 XX
 XX Sequence 306 AA;
 SQ
 Alignment Scores:
 Pred. No.: 2.54e-168 Length: 306
 Score: 1626.00 Matches: 304
 Percent Similarity: 99.35% Conservative: 0
 Best Local Similarity: 99.35% Mismatches: 2
 Query Match: 97.48% Indels: 0
 DB: 22 Gaps: 0
 US-09-735-713A-1 (1-921) x AAE04733 (1-306)
 QY 1 ATGAGTCTCAAAATGCTTATAGCAGGACAGCTGATTTTACTACTAGGATAGTCTTT 60
 Db 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuGlyIleValPhe 20
 QY 61 TTGAAACRAGGTAAATCTGCARCTTTTTCGCTCCCAAGCTCCAGTTGGGCAGAGT 120
 Db 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyClnSer 40
 QY 121 CTGGTTAAGGTACAGCTTGAATTTTAAACATTTTTCAGTCGATTTCTTGGAGGAAGC 180
 Db 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
 QY 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGCAGACAGATATT 240
 Db 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
 QY 241 TGTGAGGAAGCATGCTCTACACAGTGGGTGATCAGCGGGCTCACTGCATTGCCAAAC 300
 Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
 QY 301 AGAAACATTGCTGCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCAGACAGAC 360
 Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
 QY 361 CCAGGAGAGCAACTCTCATTGAACTGTCATCATCATCATCATCATCATCATCATCATCAT 420
 Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140

QY 421 AAACCAATGGAGTATATATTTCCCTTTTGAAGATGGTGGAGCCTTCCAAATTTGGCCAC 480
 Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
 QY 481 TTTGTGGGGCCCATATGCTTTCAGAGCTGGGGAGCAATTTGAGCGCTGGTTTATTTGT 540
 Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
 QY 541 ACAACTGCAGGCTGGGGCGGCTTAACCTGAAGTGGCGCTCTCTCACAAGTCTTGAGAA 600
 Db 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
 QY 601 GTGAATCTGCCTATTGACCTGGGAAGAGTGTGGCAGCTCTCTTAACACTAAAGAGG 660
 Db 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
 QY 661 CCCATCAGTGGAGACCTTTCTTTCACAGCTTTCTGATGGAGGAGAGACGATGT 720
 Db 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
 QY 721 CAGGAGATTGAGGAGTTCATCTATGTCGCGGATTAAGAAGGGCCCTGACTCGCT 780
 Db 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyValaTrpThrLeuAla 260
 QY 781 GGTGTGACTTCTGGGGTTTGGGCTGTGTCGAGCTGGAGAACAAATGTGAGGAAAGT 840
 Db 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280
 QY 841 GATCAAGGATCCCTGGGATCTTACAGACATTAGTAAAGTCTTCTGATCCACGAA 900
 Db 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
 QY 901 CACATCCAACTGGTAAAC 918
 Db 301 HistIleGlnThrGlyAsn 306
 RESULT 2
 AAU74760
 ID AAU74760 standard; Protein: 306 AA.
 XX
 AC AAU74760;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human protease PRTS-20 protein sequence.
 XX
 KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200198468-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001WO-US19178.
 XX
 PR 16-JUN-2000; 2000US-212336P.
 PR 22-JUN-2000; 2000US-213955P.
 PR 29-JUN-2000; 2000US-215396P.
 PR 07-JUL-2000; 2000US-216821P.
 PR 14-JUL-2000; 2000US-218946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;

PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX WPI: 2002-090437/12.
DR N-PSDB; ABK12903.
XX
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
PT proliferative (e.g. cancer) disorders -
XX
XX Claim 1; Page 157-158; 177pp; English.
XX
XX The present invention relates to twenty one new human proteases,
CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
CC myocardial infarction, autoimmune/inflammatory e.g. acquired
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
CC endometriosis disorders. Numerous other examples of each disorder are
CC given in the specification. The present protein sequence represents
CC the human protease PRTS-20 protein of the invention.
XX
XX Sequence 306 AA;

Alignment Scores:

Pred. No.: 2,54e-168 Length: 306
Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservativity: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 97.48% Indels: 0
DB: 23 Gaps: 0

US-09-735-713A-1 (1-921) x AAU74760 (1-306)

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTAGGAATAGTCTTTT 60
DB 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuLeuLeuValPhe 20
QY 61 TTGCAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 120
DB 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
QY 121 CTGTTAAGGTACAGCTTGAATATTATTAACATTTTTCAGTCGCATCTTGGAGGAGC 180
DB 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
QY 181 CAAGTGGAGAGGCTCTCTATCCCTGGCAGGTATCTCTGAACAAGCAGCATATT 240
DB 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
QY 241 TGTGGAGGAGCAGCTGCTCACACAGTGGGTGATCACGGCGGCTCACTGCATTGCAAAAC 300
DB 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
QY 301 AGAACATTTGCTACTTGAATGTTACTGCTCGAGAGATGATGACTTAAGCCAGACAGAC 360
DB 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
QY 361 CCAGGAGGCAACTCTCACTATTGAACCTCATCATCATCATCATCATCATCATCATCAT 420
DB 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
QY 421 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCTTCCCAATTTGGCCAC 480
DB 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
QY 481 TTTCTGGGGCCCATATCTCTTCCAGAGCTGCGGAGCAATTTGAGCTGGTTTATTGTT 540
|||||

DB 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
QY 541 ACAACTGCAGCTGGCGCTTAACCTGAAGTGGCGCTCTCACAGTCTTGCAGCAA 600
DB 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200
QY 601 GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
DB 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
QY 661 CCCATCAGTGGGAAGACCTTTCTTTGGCACAGGTTTTCTGTATGGAGGAGAGCGCATGT 720
DB 221 ProfileSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
QY 721 CAGGAGATTCAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGGCTGGACTCTGGCT 780
DB 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
QY 781 GGTGTGACTTCTCGGGTGTGGGTGTGGTGTGGAGCTGGAGAAACAATGTGAGGAAAAAGT 840
DB 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280
QY 841 GATCAAGGATCCCTGGGATCTTCACAGACATTTAGTAAAGTCTTCTCGATCCACGAA 900
DB 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
QY 901 CACATCCAACTGCTAAC 918
DB 301 HisIleGlnThrGlyAsn 306
RESULT 3
AAE04734
ID AAE04734 standard; Protein: 302 AA.
XX
AC AAE04734;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human protease homologue #2.
XX
KW Human; protease homologue; novel human protein; NHP; therapy;
KW pharmacogenomic application; physiological disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 19 /note= "Encoded by CRA"
FT Misc-difference 24 /note= "Encoded by RCT"
XX
XX WO200146407-A1.
XX
XX 28-JUN-2001.
XX
XX 12-DEC-2000; 2000WO-US33738.
XX
XX 23-DEC-1999; 99US-0171566.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX WPI: 2001-408641/43.
XX N-PSDB; AAD09329.
XX
XX Polynucleotide encoding novel human protease homologs, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
PS Disclosure; Page 29-30; 32pp; English.
XX

SQ Sequence 556 AA;

Alignment Scores:

Pred. No.: 9,24e-139 Length: 556
Score: 1356.00 Matches: 262
Percent Similarity: 87.95% Conservative: 8
Best Local Similarity: 85.34% Mismatches: 29
Query Match: 81.29% Indels: 8
DB: 23 Gaps: 2

US-09-735-713A-1 (1-921) x AAU82748 (1-556)

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGAACAAAGCTGATTCTTACTAGGAATAGTCTTT 60
Db 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTCGCTCCCAAGCTCCCAAGTGTGTGGCGAGAGT 120
Db 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
QY 121 CTGTTTAAGTACAGCTTGGATTATTTAAACATTTTCAGTCGCATTTCTGGAGGAAGC 180
Db 41 LeuValLysValGlnProIrpAsnIrpPheAsnIlePheSerArgIleLeuGlyGlySer 60
QY 181 CAAAGTGAAGAGGTTCTATCCCTGCAGGTATCTCTGAACAAAGGAGGAGCATATT 240
Db 61 GlnValGluLysGlySerTyrProIrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
QY 241 TGTGGAGGAAGCATCTCTACACAGTGGGTGATCAGCGCGCTCACTGCATTTGCAAAAC 300
Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
QY 301 AGAACATCTGTCTACTTGTATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360
Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
QY 361 CCAGGAGAGCAAACTCTCACTATTGAACTGTATCATCATATCCACATTTCTCCACCAAG 420
Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
QY 421 AAACCAATGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCAC 480
Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
QY 481 TTTGTGGGCCCATATGCTTCCAGAGCTCCGGAGCAATTTGAGCTGGTTTATTCTT 540
Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
QY 541 ACAACTGCAGGCTGGGGCCGCTTAAGTGAAGTGGGCTCTCTCACAAGTCTTTGCAGGAA 600
Db 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200
QY 601 GTCAATCTGCTATTTTGACCTGGGAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
Db 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
QY 661 CCCATCAGTGGGAAGACCTTCTTTCACAGGTTTCTTCATGCGAGGAGACGCATGT 720
Db 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyArgAspAlaCys 240
QY 721 CAGGAGATTCAGAGGTTCACTCATGTGCGGAATAAGAAAGGCGCTGGACTCTGGCT 780
Db 241 GlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaIrpAsp---Ser 259
QY 781 GGTGTGACTTCCGGGGTTGGGCTGTGGTGGAGCGTGGAGAAACAATGTGAGGAAAGT 840
Db 260 GlyTrpSerIleTrpGluAlaGlnValGlySerLeuGluSerArgSerArgPro 279
QY 841 GATCAAGGATCCCT-----GGGATCTTCACAGACATTAGTAA 879
Db 280 SerLeuGlyAsnLysValArgLeuCysLeuThrAsnAsnPhePheLysLysLeuAlaGly 299
QY 880 GTGCTTTTCTGGATCCACGAA 900

Db 300 CysGlyThrTrpCysSerGlu 306
RESULT 5
AAM24275

ID AAM24275 standard; Protein; 182 AA.
XX
AC AAM24275;

XX
DT 12-OCT-2001 (first entry)
XX
DE Toad EST encoded protein SEQ ID NO: 1800.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
XX biodiversity; gene therapy; nutrition.
XX Xenopus laevis.

XX WO200154477-A2.
XX
PD 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02687.
XX
XX 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX WPI; 2001-476164/51.
DR N-PSDB; AAH98934.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
PS Claim 20; Page 1178-1179; 1275pp; English.
XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.

XX
SQ Sequence 182 AA;
Alignment Scores:
Pred. No.: 5,22e-94 Length: 182
Score: 944.00 Matches: 180
Percent Similarity: 98.90% Conservative: 0
Best Local Similarity: 98.90% Mismatches: 2
Query Match: 56.59% Indels: 0
DB: 22 Gaps: 0

US-09-735-713A-1 (1-921) x AAM24275 (1-182)
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGAACAAAGCTGATTCTTACTAGGAATAGTCTTT 60
Db 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20

QY 61 TTTGAACRAGGTAAATCTGCARCTCTTCGCTCCCAAGCTCCCAAGTGTGTGGCGAGAGT 120
Db 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40

QY 121 CTGGTTAGGTACACCCCTGGAAATATTTTAACTATTTCACTGCGATTCTTGGAGGAAGC 180
 Db 41 LeuVallysalGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer 60
 QY 181 CAAAGTGAAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGACAGACATATT 240
 Db 61 GlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
 QY 241 TGTGAGGAAGCATCGTCTCACACAGTGGGTGATCACGGCGGCTCACTGCATTCGCAAAAC 300
 Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
 QY 301 AGAAACATGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360
 Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAsp 120
 QY 361 CCAGGAGAGCAACTCTCACATTTCAAACTGCTCATCATCATCATCATCATCATCATCAT 420
 Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
 QY 421 AAACCAATGGACTATGATATTCCTCTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 480
 Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
 QY 481 TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGTTTATTGT 540
 Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
 QY 541 ACAACT 546
 Db 181 ThrThr 182
 RESULT 6
 AAE04735
 ID AAE04735 standard; Protein; 164 AA.
 AC AAE04735;
 XX
 DT 10-SEP-2001 (first entry)
 DE Human protease homologue #3.
 XX
 KW Human; protease homologue; novel human protein; NHP; therapy;
 KW pharmacogenomic application; physiological disorder.
 OS Homo sapiens.
 XX
 PN WO200146407-A1.
 XX
 PD 28-JUN-2001.
 XX
 PF 12-DEC-2000; 2000WO-US33738.
 XX
 PR 23-DEC-1999; 99US-0171566.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
 XX
 DR WPI; 2001-408641/43.
 XX
 DR N-ESDB; AAD09330.
 XX
 PT Polynucleotide encoding novel human protease homologs, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications -
 XX
 PS Disclosure; Page 30-31; 32pp; English.
 XX
 CC The present sequence is novel human protein (NHP),
 CC known as human protease homologue. NHP shares structural similarity
 CC with animal proteases, particularly trypsin-like protease such
 CC as oviductin, plasminogen activator and human plasma kallikrein

CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
 CC pharmacogenomic applications. NHP sequences are useful for identifying
 CC agonists, antagonists and modulators and also for producing antibodies
 CC useful in diagnosis, drug screening, clinical trial monitoring and in
 CC treatment of physiological disorders.
 XX
 SQ Sequence 164 AA;
 Alignment Scores:
 Pred. No.: 5,22e-89 Length: 164
 Score: 898.00 Matches: 164
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.84% Indels: 0
 DB: 22 Gaps: 0
 US-09-735-713A-1 (1-921) x AAE04735 (1-164)
 QY 427 ATGACTATGATATTCGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCACCTTTGTG 486
 Db 1 MetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 20
 QY 487 GGGCCCATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGTACAACT 546
 Db 21 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 40
 QY 547 GCAGCTGGGGCCGCTTAACCTGAAGTGGCTGCTCTCACAAAGTCTTCAGAGAGTGAAT 606
 Db 41 AlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGluValAsn 60
 QY 607 CTGCTATTTTACCTGGGAAGAGTGTGTGCAGCTCTGTTAACACTAAAGAGGCCCATC 666
 Db 61 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 80
 QY 667 ACTGGGAAGACCTTTCTTTGACAGGTTTCTGATGGAGGAGGAGCGCATGTCAGGA 726
 Db .81 SerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGly 100
 QY 727 GATTGAGGAGTTCACATCATGTCCGGAATAAGAAAGGGCTGAGCTCTGGCTGGTGTG 786
 Db 101 AspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaTrpThrLeuAlaGlyVal 120
 QY 787 ACTTCCTGGGCTTTGGGCTGTGTCGAGGCTGGAGAAACAATGAGGAAAGTGATCAA 846
 Db 121 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln 140
 QY 847 GGATCCCTGGGATCTTCACAGACATTTAGTAAAGTGTCTTCTCGATCCAGACACATC 906
 Db 141 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 160
 QY 907 CAAACTGGTAAC 918
 Db 161 GlnThrGlyAsn 164
 RESULT 7
 ABG24246
 ID ABG24246 standard; Protein; 913 AA.
 XX
 AC ABG24246;
 XX
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #24237.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX


```
PF 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS88433.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 54605; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as 'expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 913 AA;
SQ
Alignment Scores:
Pred. No.: 1,19e-52 Length: 913
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.02% Indels: 11
DB: Gaps: 4
US-09-735-713A-1 (1-921) x ABG24246 (1-913)
QY 109 TGTGGGCAGAGTCTGGTTAAGGTA-----CAGCCTTGGATATTTTAACTTTTC 159
Db 9 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 28
QY 160 AGTCGCATTCTTGGAGGACCCAGTGGAGAGGGTTCCTATCCCTGGCAGGATCTCTG 219
Db 29 SerArgIleSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 48
QY 220 AAACAAGGCGAAGCATTTTGGAGGAAGCATCTCTCCACACAGTGGGTGATCACG 279
Db 49 LysSerAspGluHisHisPheCysGlyGlySerLeuIleGlnAspArgValValThr 68
QY 280 GCGGCTCAGTCGATT-----GCAACACAGAAACATTGTCTACTTTGAATGTTACGCT 333
Db 69 AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 88
QY 334 GGAGAGTATCACTTAAGCCAGACAGACCCAGGAGGACAACTCTCACTATTGAACGTGC 393
Db 89 GlyGluTyrSerLeuPheGlnLysAspLysGlnGlnAsnIleProValSerLysIle 108
QY 394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGACATATGATATTTGCCCTTTTGAAG 453
Db 109 IleThrHisProGluTyrAsnSerArgGluTyrMetSerProAspIleAlaLeuLeuTyr 128
QY 454 ATGGCTGGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATCTCTTCCAGAGTCGCG 513
Db 129 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 148
QY 514 GAGCAATTTGAGGCTGGTTTATTTGTACAACTGCAGGCTGGGGCCCTTAACTGAAGGT 573
Db 149 AspLysValGluProGlyIleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 168
QY 574 GCGCTCTCTCAACAAGTCTTTCAGGAGTGAATCTGCCTATTTTGTACCTGGGAAGAGTGT 633
Db 169 SerGluTyrSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 188
QY 634 GTGGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTCTTTTCACAGGT 693
Db 189 AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 208
QY 694 TTTCTCTGATGGAGGAGAGAGCATGTCAGGGAGATTCAGGAGGTTCACATCATGTGCCGG 753
Db 209 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 228
QY 754 AATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCTCTGGGTGGGTGGGTGCTGA 813
Db 229 ArgGlyGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 248
QY 814 GGC-----TGGAGAAACAATGTGAGGAAAGTGTCAAGAGTCCCTGGGATCTTC 864
Db 249 GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 265
QY 865 ACAGACATTAGTAAGTGTCTTCTCGATCCACGACACATCCAAACTGCT 915
Db 266 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 282
RESULT 8
ABG07870
ID ABG07870 standard; Protein; 1576 AA.
XX
XX AC ABG07870;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #7861.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX PA Drmanac RT, Liu C, Tang YT;
XX
XX PI WPI; 2001-639362/73.
XX
XX DR N-PSDB; AAS72057.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX
XX PS Claim 20; SEQ ID No 38229; 103pp; English.
```

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG00010-ABG30377 represent novel human CC diagnostic amino acid sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1576 AA;

Alignment Scores:
 Pred. No.: 1.48e-52 Length: 1576
 Score: 567.50 Matches: 114
 Percent Similarity: 59.57% Conservative: 51
 Best Local Similarity: 41.16% Mismatches: 101
 Query Match: 34.02% Indels: 11
 DB: 22 Gaps: 4

US-09-735-713A-1 (1-921) x ABG07870 (1-1576)

QY 109 TGTGGCAGAGCTGGTTAAGTA-----CAGCCTTGGAAATATTAAACATTTTC 159
 DB 848 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 867
 QY 160 AGTCGCATCTTGGAGGAGCCAGTGGAGAGGGTTCTCTCCCTGGCAGGTATCTCTG 219
 DB 868 SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 887
 QY 220 AAACAAGGAGAGCATATTGTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGC 279
 DB 888 LysSerAspGluHisPheCysGlySerLeuIleGlnGluAspArgValValThr 907
 QY 280 GCGGCTCACTGAT------GCAACAGAACATTTGTCTACTTTGAATGTACTGCT 333
 DB 908 AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 927
 QY 334 GGAGAGTATGACTTAAGCCAGACAGCCAGGAGAGCAAACTCTCAGTATTGAAACTGTC 393
 DB 928 GlyGluTrpSerLeuPheGlnLysAspLysGlnGluGlnAsnIleProValSerLysIle 947
 QY 394 ATCAPACATCCACATTTCTCCACCAAGAACCAATGGCATATGATTCGCCCTTTTGAAG 453
 DB 948 IleThrHisProGluLysAsnSerArgGluTrpMetSerProAspIleAlaLeuLysTrp 967
 QY 454 ATGGCTGGAGCCTCCATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTGGG 513
 DB 968 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987
 QY 514 GAGCAATTTAGGCTGGTTATTGTACAACTGCAGGCTGGGGCGCTTAAGTGAAGT 573
 DB 988 AspLysValGluProGlyIleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 1007
 QY 574 GCGGCTCTCCACAGTCTTCAGGAAGTGAATTCGCTATTTTGAACCTGGGAAGAGTGT 633
 DB 1008 SerGluTrpSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 1027
 QY 634 GTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACAGGT 693

Db 1028 AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 1047
 QY 694 TTTCTGTATGGAGAGAGAGCATGTCTCAGGAGATTTCAGGAGTTCATCTCATGTCCGG 753
 Db 1048 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 1067
 QY 754 AATAAGAAAGGGCCCTGGACTCTGGCTGTGTGACTTCTCTGGGGTTTGGGTGGTCTCGA 813
 Db 1068 ArgGlyGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 1087
 QY 814 GGC-----TGGAGAAACAATGTGAGGAAAGTATCAAGGATCCCTGGGATCTTC 864
 Db 1088 GlySerValProValArgAsnHisValLysAla-----SerLeuGlyIlePhe 1104
 QY 865 ACAGACATTAGTAAGTCTTCTTCGTGATCCAGCAACATCCAACTGGT 915
 Db 1105 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 1121
 RESULT 9
 ABG10218
 ID ABG10218 standard; Protein: 1576 AA.
 XX
 AC ABG10218;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #10209.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74405.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 40577; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX	Sequence	1576 AA;
SQ	Alignment Scores:	
	Pred. No.:	1.48e-52
	Score:	567.50
	Percent Similarity:	59.57%
	Best Local Similarity:	41.16%
	Query Match:	34.02%
	Indels:	11
	Gaps:	4
DB:		
US-09-735-713A-1 (1-921) x ABG10218 (1-1576)		
QY	109	TGTGGCAGAGTCTGTTAAGTA-----CAGCCTTGGAAATATTTTAACATTTTC 159
DB	848	CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 867
QY	160	AGTCGATCTCTGGAGGAGCCCAAGTGGAGAGGGTTCCTCCCTGGCAGGATCTCTCG 219
DB	868	SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 887
QY	220	AAACAAGCAGAGCATATTTCTGGGAGGAGCATGCTCCACACAGTGGGTGATCAGC 279
DB	888	LysSerAspGluHisPheCysGlySerLysGluLeuGlnGluAspArgValValThr 907
QY	280	GCGGCTCACTGCATT-----GCAACACAGAAACATTGTGTCTACTTTGAATGTACTGCT 333
DB	908	AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 927
QY	334	GGAGATATGACTTAGCCAGCAGACACCAGGAGCAACTCTCACTATTGAACTGTC 393
DB	928	GlyGluTrpSerLeuPheGlnLysAspLysGlnGluAsnIleProValSerLysIle 947
QY	394	ATCATACATCCACATTTCTCCACCAAGAAACCAATGATATATGTCCTCTTTTGAAG 453
DB	948	IleThrHisProGluTrpAsnSerArgGluTrpMetSerProAspIleAlaLeuLeuTrp 967
QY	454	ATGGCTGGAGCCCTCCAAATTTGGCCACTTTTGGGGCCCATATGCTTCCAGAGCTGCGG 513
DB	968	LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987
QY	514	GACCAATTTGAGGCTGGTTTATTTTACAACTGCGAGGTGGGGCCGCTTAAGTGAAGT 573
DB	988	AspLysValGluProGlyIleLeuLysLeuSerSerGlyTrpGlyLysIleSerLysThr 1007
QY	574	GCGCTCTCTCACAAGTCTTGAGGAAGTGAATCTGCTATTTTGACCTGGGAAGAGTGT 633
DB	1008	SerGluTrpSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 1027
QY	634	GTGGCAGCTCTGTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTGACAGGT 693
DB	1028	AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 1047
QY	694	TTTCTCTGATGGAGGAGACGATCTCAGGAGATTCAGGAGGTTCATCTATGTCGCGG 753
DB	1048	PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 1067
QY	754	AATAAGAAAGGGCCCTGGACTCTGGCTGTGTGCTGCTCTCTCTCTCTCTCTCTCTCGA 813
DB	1068	ArgGlyGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 1087
QY	814	GGC-----TGGAGAAACAATGTGAGGAAAGATGATCAAGGATCCCTGGGATCTTC 864
DB	1088	GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 1104
QY	865	ACAGACATAGTAAAGTCTTCTCTGATCCACGACACATCCAAACTGCT 915
DB	1105	SerLysValSerGluLeuMetAspPheIleThrGluAsnLeuPheThrGly 1121

RESULT 10	
ABG14588	
ID	ABG14588 standard; Protein; 1576 AA.
XX	
AC	ABG14588;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #14579.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO2001/5067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS78775.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
PT	
PT	
XX	
PS	Claim 20; SEQ ID No 44947; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 1576 AA;
Alignment Scores:	
Pred. No.:	1.48e-52
Score:	567.50
Percent Similarity:	59.57%
Best Local Similarity:	41.16%
Query Match:	34.02%
Indels:	11
Gaps:	4
DB:	
US-09-735-713A-1 (1-921) x ABG14588 (1-1576)	
QY	109 TGTGGCAGAGTCTGGTTAAGTA-----CAGCCTTGGAAATATTTTAACATTTTC 159

```
Db 848 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 867
QY 160 AGTCGCATTCTGGAGGAGCAAGTGGAGAGGTTCTCTATCCTCGCAGGTATCTCTG 219
Db 868 SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 887
QY 220 AAACAAGGACAGACATATTGTGGAGGAGCATCTCTCACCACAGTGGGTGATCAG 279
Db 888 LysSerAspGluHisHisPheCysGlySerLeuIleGlnGluAspArgValValThr 907
QY 280 GCGGCTCACTGCATT-----GCAACAGAAACATTTGCTTCTACTTGAATGTACTGCT 333
Db 908 AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 927
QY 334 GGAGAGTAGTAAAGCCAGACAGACCCAGGAGAGCAAACTCTCACATATTGAAACTGTC 393
Db 928 GlyGluTyrSerLeuPheGlnLysAspLysGlnGluGlnAsnIleProValSerLysIle 947
QY 394 ATCATATCATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTTGAAG 453
Db 948 IleThrHisProGluTyrAsnSerArgGluTyrMetSerProAspIleAlaLeuTyr 967
QY 454 ATGGCTGAGGCTTCCAAATTGGCCACTTTGGGGCCCATATGTCTTCCAGAGCTGGG 513
Db 968 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987
QY 514 GAGCAATTTAGGCTGTTTATTGTGTACACTGAGCTGGGGCCGCTTAACCTGAAGGT 573
Db 988 AspLysValGluProGlyLeuLysSerGlySerGlyTrpGlyLysIleSerLysThr 1007
QY 574 GGCGTCTCTCACAGTCTTCAGGAGTGAATCTGCCTATTTTACCTGGGAGAGTGT 633
Db 1008 SerGluTyrSerAsnValLeuGlnMetGluLeuProIleMetAspArgAlaCys 1027
QY 634 GTGGCAGCTCTTTAACTAAAGGCCCATCATAGTGGGAGACCTTTCTTTGGCAGAGT 693
Db 1028 AsnThrValLeuLysSerMetAsnLeuProLeuGlyArgThrMetLeuCysAlaGly 1047
QY 694 TTTCTGTATGGAGGAGAGAGCATGTCTCAGGAGATTCAGGAGGTCTCACTCATGTGCGG 753
Db 1048 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 1067
QY 754 AATAAGAAAGGGCTGACTCTGGCTGTGTGACTCTCTGGGTGGGTGGGTGGTGGTGA 813
Db 1068 ArgGlyGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 1087
QY 814 GGC-----TGGAGAAACAATGTGAGGAAAGTGATCAAGGATCCCTGGGATCTTC 864
Db 1088 GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 1104
QY 865 ACAGACATTAGTAAGTCTTCTTCTGATCCACGAAACACATCCAAACTGGT 915
Db 1105 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 1121
RESULT 11
ABG19887
ID ABG19887 standard; Protein; 1576 AA.
AC
AC ABG19887;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19878.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
```

```
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS84074.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID NO 50246; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of DNA and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1576 AA;
Alignment Scores:
Pred. No.: 1.48e-52 Length: 1576
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.02% Indels: 11
DB: 22 Gaps: 4
US-09-735-713A-1 (1-921) x ABG19887 (1-1576)
QY 109 TGTGGCAGAGCTGTGGTTAAGGTA-----CAGCCTTGGAAATATTTTAAACATTTTC 159
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 848 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 867
QY 160 AGTCGCATTCTTGGAGGAGCAAGTGGAGAGGTTCTCTATCCTCGCAGGTATCTCTG 219
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 868 SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 887
QY 220 AAACAAGGACAGACATATTGTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGC 279
| | | | | : : : : : | | | | | : : : : : : : : : : : : : : : :
Db 888 LysSerAspGluHisHisPheCysGlySerLeuIleGlnGluAspArgValValThr 907
QY 280 GCGGCTCACTGCATT-----GCAACAGAAACATTTGCTTCTACTTGAATGTACTGCT 333
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 908 AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 927
QY 334 GGAGAGTAGTAAAGCCAGACAGACCCAGGAGAGCAAACTCTCACATATTGAAACTGTC 393
| | | | | : : : : : | | | | | : : : : : : : : : : : : : : : :
Db 928 GlyGluTyrSerLeuPheGlnLysAspLysGlnGluGlnAsnIleProValSerLysIle 947
```

```
QY 394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTTGAAG 453
   ||| ||||| : : : : : : : : : ||| ||||| ||||| |||||
Db 948 lIeThrHisProGluTyAsnSerArgGluTyMetSerProAspIleAlaLeuLeuTy 967
QY 454 ATGGCTGGAGCCTTCCAAATTTGGCCACATTTGGGGCCCATATGCTCCAGAGCTGGG 513
   : : : : : : : : : ||| ||||| ||||| ||||| : : :
Db 968 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987
QY 514 GAGCAATTTGAGCGCTGGTTTATTTGTACAACCTGCAGGCTGGCGCTTAACCTGAAGT 573
   : : : : : ||| ||| : : : : : : : : : : : : : : : : : :
Db 988 AspLysValGluProGlyIleLeuCysLeuSerGlyTrpGlyLysIleSerLysThr 1007
QY 574 GCGCTCTCTCACAAGCTCTGCGAGGAAGTGAATCTGCCTATTGACCTGGGAAGAGTGT 633
   ||| ||||| ||||| ||||| : : : : : : : : : : : : : :
Db 1008 SerGluTyTrSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 1027
QY 634 GTGGCAGCTCTGTTAACTAAGAGGCCCATCAGTGGGAAGACCTTTCTTGCACAGT 693
   : : : : : : : : : ||| ||||| ||||| ||||| |||||
Db 1028 AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 1047
QY 694 TTTCTGTGATGAGGAGACACATGTCAGGAGATTTCAGGAGTTTCACCTCATGTGCGCG 753
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1048 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 1067
QY 754 AATAAGAAAGGGCGCTGACTGCTGGTGTGCTGCTCTCTGGGGTTTGGCTCTGCTCGA 813
   ||| ||| ||||| ||||| : : : : : : : : : |||||
Db 1068 ArgGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 1087
QY 814 GGC-----TGAGAAACAATGTGAGGAAGATGATCAAGATGCCCTGGGATCTTC 864
   ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1088 GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 1104
QY 865 ACAGACATTAGTAAGTGTCTTCTCGATCCACAGACATCAAACTGCT 915
   : : : : : : : : : : : : : : : : : : : : : : :
Db 1105 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 1121

RESULT 12
ID ABP60993
XX ABP60993 standard; Protein; 1031 AA.
AC ABP60993;
XX
XX
DT 10-SEP-2002 (first entry)
DE
DE
XX Novel human protein. SEQ ID 80.
XX
KW Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiatic; antiulcer; virucide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
XX
OS Homo sapiens.
XX
PN WO200250105-A1.
XX
PD
PD
XX
XX
PF 17-DEC-2001; 2001WO-US49232.
XX
XX
PR 19-DEC-2000; 2000US-256710P.
```

```
PR 20-DEC-2000; 2000US-257048P.
PR 09-JAN-2001; 2001US-260482P.
PR 30-JAN-2001; 2001US-264922P.
PR 06-FEB-2001; 2001US-266797P.
PR 19-MAR-2001; 2001US-276988P.
PR 04-APR-2001; 2001US-281535P.
PR 08-MAY-2001; 2001US-289622P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX (GLAX ) GLAXO GROUP LTD.
PI
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX WPI; 2002-508784/54.
XX N-PSDB; ABQ86158.
PT
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune
PT disorder -
XX
XX Claim 1(a); Page 299-301; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
XX which allow it to be secreted extracellularly or membrane associated.
XX The activity of polypeptides of the invention may be described as,
XX cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, nootropic,
XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX cardiatic, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,
XX and metabolic. Polypeptides and polynucleotides of the invention are
XX useful in the treatment, or as a vaccine in the prevention of, cancer,
XX wound healing disorders, infection, atherosclerosis, Parkinson's disease,
XX and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX inflammation, neoplastic diseases, nervous system related disorders and
XX cardiovascular disorders, pancreatitis, respiratory disorder,
XX hyperproliferation, systemic autoimmune disease, hyper-immunity,
XX developmental abnormality, gastrointestinal ulceration, neuropathy,
XX haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
XX transduction deficiency, neurological diseases, stroke, anglogenesis,
XX ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
XX trachea, thymus, lymph node and muscular system, obesity, anorexia,
XX growth abnormalities, and alleviation of precocious puberty. The
XX sequences given in records ABP60965-ABP61019 represent novel human
XX proteins of the invention.
```

Sequence 1031 AA;

Alignment Scores:
Pred. No.: 5,62e-52 Length: 1031
Score: 561.50 Matches: 113
Percent Similarity: 59.42% Conservative: 51
Best Local Similarity: 40.94% Mismatches: 101
Query Match: 33.66% Indels: 11
DB: 23 Gaps: 4

US-09-735-713A-1 (1-921) x ABP60993 (1-1031)

```
QY 109 TGTGGCAGAGTCTGGTTAAGTA-----CAGCCTTGGATTTATTTAACATTTTC 159
   ||||| : : : : : : : : : |||||
Db 9 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 28
QY 160 AGTCGATCTTGGAGGAGCCCAAGTGGAGAGGGTTCCTATCCCTGGCAGGATCTCTG 219
   ||||| ||||| : : : : : : : : : |||||
Db 29 SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 48
QY 220 AAACAAGGCACAGCATATTTGTGAGGAACATCGTCTCACCACAGTGGGTGATCAG 279
   ||||| : : : ||| ||||| ||||| : : : : : : : : : |||||
Db 49 LysSerAspGluHisPheCysGlyGlySerLeuIleGlnGluAspArgValThr 68
QY 280 GCGGCTCACTGCATT-----GCAACAGAAACATTTGTCTACTTTGAATGTTACTGCT 333
```

```
Db      69  AlalaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 88
      |||||
Qy      334  GGAGCATGATGCTTAAGCCAGACAGACCCAGGAGAGAACTCTCACTATTGAACCTGTC 393
      |||||
Db      89  GlyGluThrSerLeuPheGlnLysAspLysGlnGluGlnAsnIleProValSerLysIle 108
      |||||
Qy      394  ATCATACATCACATTTCTCCACCAAGAACCAATGACATGATATTTGCCCTTTTGAAG 453
      |||||
Db      109  IleThrHisProGluThrAsnSerArgGluThrMetSerProAspIleAlaLeuLeuThr 128
      |||||
Qy      454  ATGCTCGAGCGCTTCCAAATTTGCCACTTTGTGGGGCCCATATGCTTCTCCAGAGCTCGG 513
      |||||
Db      129  LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 148
      |||||
Qy      514  GAGCAATTTGAGGCTGCTTTTATTGTAACATCCAGCTGGGGCCCTTAACCTGAAGGT 573
      |||||
Db      149  AspLysValGluProGlyIleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 168
      |||||
Qy      574  GGCCTCCTCTCACAAAGTCTTGAGAAAGTGAATCTGCTATTGACCTGGGAAGAGTGT 633
      |||||
Db      169  SerGluThrSerAsnValLeuGlnLuwMetGluLeuProIleMetAspArgAlaCys 188
      |||||
Qy      634  GTGCAGCTGCTGTAAACACTAAAGAGCCCATCATGCTGGGAAGACCTTCTTTCACACAGT 693
      |||||
Db      189  AsnThrValLeuLysSerMetAsnLeuProLeuGlyArgThrMetLeuCysAlaGly 208
      |||||
Qy      694  TTTCTCATGAGGAGGAGGAGCGCATGTCAGGAGATTCAGAGGTTTCACTCATGTCGCGG 753
      |||||
Db      209  PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 228
      |||||
Qy      754  AATAAGAAAGGGGCGCTGACTCTGGCTGGTGTGACTTCCCTGGGCTTGGGCTGTGGTCGA 813
      |||||
Db      229  ArgGlyGlyGlyIleThrPheLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 248
      |||||
Qy      814  GGC-----TGGAGAACATGTGAGGAAAGATGATCAAGGATCCCTGGGATCTTC 864
      |||||
Db      249  GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 265
      |||||
Qy      865  ACACACATTTAGTAAGTCTTCTCGTGCATCCACGACACATCCAAACT 912
      |||||
Db      266  SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThr 281
      |||||

RESULT 13
AAY41710
ID      AAY41710 standard; Protein; 802 AA.
XX
AC      AAY41710;
XX
DT      07-DEC-1999 (first entry)
XX
DE      Human PRO618 protein sequence.
XX
KW      Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW      probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW      secreted protein; transmembrane protein.
XX
OS      Homo sapiens.
XX
PN      W09946281-A2.
XX
PD      16-SEP-1999.
XX
PF      08-MAR-1999; 99WO-US05028.
XX
PR      10-MAR-1998; 98US-0077450.
PR      11-MAR-1998; 98US-0077632.
PR      11-MAR-1998; 98US-0077641.
PR      11-MAR-1998; 98US-0077649.
PR      12-MAR-1998; 98US-0077791.
PR      13-MAR-1998; 98US-0078004.
PR      17-MAR-1998; 98US-0040220.
PR      20-MAR-1998; 98US-0078886.
PR      20-MAR-1998; 98US-0078910.
PR      20-MAR-1998; 98US-0078936.
PR      20-MAR-1998; 98US-0078939.
PR      25-MAR-1998; 98US-0079294.
PR      26-MAR-1998; 98US-0079656.
PR      27-MAR-1998; 98US-0079663.
PR      27-MAR-1998; 98US-0079664.
PR      27-MAR-1998; 98US-0079689.
PR      27-MAR-1998; 98US-0079728.
PR      27-MAR-1998; 98US-0079786.
PR      30-MAR-1998; 98US-0079920.
PR      31-MAR-1998; 98US-0079923.
PR      31-MAR-1998; 98US-0080105.
PR      31-MAR-1998; 98US-0080107.
PR      31-MAR-1998; 98US-0080165.
PR      31-MAR-1998; 98US-0080194.
PR      01-APR-1998; 98US-0080327.
PR      01-APR-1998; 98US-0080328.
PR      01-APR-1998; 98US-0080333.
PR      01-APR-1998; 98US-0080334.
PR      08-APR-1998; 98US-0080334.
PR      08-APR-1998; 98US-0081049.
PR      08-APR-1998; 98US-0081070.
PR      09-APR-1998; 98US-0081195.
PR      09-APR-1998; 98US-0081203.
PR      09-APR-1998; 98US-0081229.
PR      15-APR-1998; 98US-0081817.
PR      15-APR-1998; 98US-0081838.
PR      15-APR-1998; 98US-0081952.
PR      15-APR-1998; 98US-0081955.
PR      21-APR-1998; 98US-0082568.
PR      21-APR-1998; 98US-0082569.
PR      22-APR-1998; 98US-0082700.
PR      22-APR-1998; 98US-0082704.
PR      22-APR-1998; 98US-0082804.
PR      23-APR-1998; 98US-0082767.
PR      23-APR-1998; 98US-0082796.
PR      27-APR-1998; 98US-0083336.
PR      28-APR-1998; 98US-0083322.
PR      29-APR-1998; 98US-0083392.
PR      29-APR-1998; 98US-0083495.
PR      29-APR-1998; 98US-0083496.
PR      29-APR-1998; 98US-0083499.
PR      29-APR-1998; 98US-0083500.
PR      29-APR-1998; 98US-0083545.
PR      29-APR-1998; 98US-0083554.
PR      29-APR-1998; 98US-0083558.
PR      29-APR-1998; 98US-0083559.
PR      30-APR-1998; 98US-0083742.
PR      05-MAY-1998; 98US-0084366.
PR      06-MAY-1998; 98US-0084414.
PR      06-MAY-1998; 98US-0084441.
PR      07-MAY-1998; 98US-0084598.
PR      07-MAY-1998; 98US-0084600.
PR      07-MAY-1998; 98US-0084627.
PR      07-MAY-1998; 98US-0084637.
PR      07-MAY-1998; 98US-0084639.
PR      07-MAY-1998; 98US-0084640.
PR      07-MAY-1998; 98US-0084643.
PR      13-MAY-1998; 98US-0085323.
PR      13-MAY-1998; 98US-0085338.
PR      13-MAY-1998; 98US-0085339.
PR      15-MAY-1998; 98US-0085573.
PR      15-MAY-1998; 98US-0085579.
PR      15-MAY-1998; 98US-0085580.
PR      15-MAY-1998; 98US-0085582.
PR      15-MAY-1998; 98US-0085689.
PR      15-MAY-1998; 98US-0085697.
PR      15-MAY-1998; 98US-0085700.
PR      15-MAY-1998; 98US-0085704.
PR      18-MAY-1998; 98US-0086023.
PR      22-MAY-1998; 98US-0086392.
PR      22-MAY-1998; 98US-0086414.
```

PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH) GENENTECH INC.
XX
XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
PI N-PSDB; AAZ34033.
XX
XX WPI; 1999-551358/46.
DR N-PSDB; AAZ34033.
XX
XX New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders .
XX
XX Claim 12; Fig 63; 530pp; English.
XX
XX The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AAZ33891 to
CC AAZ34338, and AA41685 to AA41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.
XX
XX SQ Sequence 802 AA;

Alignment Scores:
Pred. No.: 3,648-44 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 20 Gaps: 5

US-09-735-713A-1 (1-921) x AA41710 (1-802)

QY 160 ACTCGATTCTGGAGGAGCCAGTGGAGAGGGTTCTATCCCTGGCAGGTATCTGTG 219
DB SerArgileValGlyGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585

QY 220 AAACAAAGGAGAGCATATTTGTGGAGGAGCATCTCTCCACAGTGGGTGATCAGC 279
DB GlnValArgGlyArgHisIleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThr 605

QY 280 GGGCTCACTGATTCACAAACAAACATGTGTCTACTTTG-----AATGTTACTGCT 333
DB AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrPheLeu 625

QY 334 GGAGACTATGACTTAAGCCAGACACCCAGGAGACCAACTCTCACTATTGAAACTGTC 393
DB GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645

QY 394 ATCATATCCACATTTCTCCCAAGAACCAATGGAGCTATGATATTCCTTGTGAAG 453
DB LeuLeuHisPro--TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664

QY 454 ATGGCTGGAGCCCTCCAAATTTGGCCACTTTGGGGGCCCATATGTCTCCAGAGTGGG 513
DB LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684

QY 514 GACCAATTTGAGGCTGTTTATTTTCTACAACTGCAGGCTGGGGCGCTTAACGTGAAGT 573
DB HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704

QY 574 GGCCTCTCTCACAGTCTTGCAGGAAGTGAATCTGCCTATTTTACCTGGGAAGAGTGT 633
DB GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724

QY 634 GTGGCAGCTCTGTTAACTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTTGCACAGT 693
DB SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740

QY 694 TTTCTGATGGAGGAGAGACGATCTCAGGAGAGATTAGGAGGTTCATCATGTGCCGG 753
DB TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760

QY 754 AATAAGAAAGGGCCTGACTCTGCTGCTGTGTGACTTCTCTGGGTTGGCTGTGTGTCGA 813
DB AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780

QY 814 GCCTGGAGAAACAAATGTGAGGAAAGTATCAAGGATCCCT-----GGGATCTTC 864
DB 781 -----ProAsnTyrPheGlyValTyr 787

QY 865 ACAGACATTAGTAAAGTCTTTCTGATCCACGAA 900
DB ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 14
AAB44266
ID AAB44266 standard; Protein; 802 AA.
XX
AC AAB44266;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO618 (UNQ354) protein sequence SEQ ID NO:169.
KW Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
KW expressed sequence tag; detection; cancer.
OS Homo sapiens.
XX
PN WO200053756-A2.
XX
PD 14-SEP-2000.
XX
PF 18-FEB-2000; 2000WO-US04341.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
DR WPI; 2000-611443/58.

27

DR N-PSDB; AAC78494.
XX Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities -
XX
XX
XX Claim 12; Fig 63; 636pp; English.
XX
CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting
CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX
SQ Sequence 802 AA;

Alignment Scores:
Pred. No.: 3.64e-44 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 21 Gaps: 5

US-09-735-713A-1 (1-921) x AAB44266 (1-802)
QY 160 AGTCGATCTTGGAGGAGCAAGTGGAGAGGTTCTATCCCTGCGCAGGTATCTCTG 219
DB 566 SerArgIleValGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGAGCAAGCATATTGTGGAGAGCATGCTCTCACCACAGTGGGTGATCAGC 279
DB 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGCTCACTGCATGCAACAGAAACATTTGCTACTTTG-----AATGTTACTGCT 333
DB 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTAGTACTTAAGCAGACAGACCCAGGAGCAAACTCTCACTATTGAAACGTGC 393
DB 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATGCTCTTTTGAAG 453
DB 646 LeuLeuHisPro--TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664
QY 454 ATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTGGG 513
DB 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GAGCAATTTAGGCTGGTTTATTGTTACAACTCAGCTGGCGCCGCTTAACCTGAAGGT 573
DB 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
QY 574 GCGCTCTCTCACAGTCTTCAGGAAGTGAATCTGCTATTTTGAAGCTGGGAAGAGTGT 633
DB 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLysCys 724
QY 634 GTGGCAGCTCTGTTAACTAAAGAGGCCCATCATAGTGGGAAGACCTTTCTTTGACAGGT 693
DB 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCTGTATGAGGAGAGCGCATGTCAGGAGATTCAGAGCTTCACTCATGTGCGGG 753
DB 741 TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGGCCTGGACTCTGGCTGGTGGTGTGACTTCTCTGGGTTTGGGTGTGTCGA 813

DB 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GCCTGGAGAAACAATGTGAGGAAAGTCAAGATCCCT-----GGGATCTTC 864
DB 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATAGTAAGTCTTCTTCTGATCCACGAA 900
DB 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799
RESULT 15
AAB24052
ID AAB24052 standard; Protein; 802 AA.
XX AAB24052;
AC AAB24052;
XX
DT 25-JAN-2001 (first entry)
XX
DE Human PRO618 protein sequence SEQ ID NO:24.
XX
KW Human; tumour; diagnosis; neoplastic disease; identification; cancer;
KW tumorigenesis; detection; neoplastic cell growth; proliferation;
KW cytostatic; antiinflammatory; immunomodulatory; inflammatory disorder;
KW immunological disorder.
OS Homo sapiens.
XX
PN WO200053754-A1.
XX
PD 14-SEP-2000.
XX
PF 06-JAN-2000; 2000WO-US00277.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 05-OCT-1999; 99WO-US23089.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US28564.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desauvage FJ, Goddard A, Gurney AL, Klein RD, Roy MA;
PI Wood WI;
XX
XX WPI; 2000-572269/53.
DR N-PSDB; AAC58236.
XX
XX New isolated antibody for use in compositions and methods for the
PT diagnosis and treatment of neoplastic cell growth and proliferation in
PT mammals, including humans, and in monitoring tumor treatment -
XX
PS Claim 61; Fig 24; 195pp; English.
XX
XX The present invention describes an isolated antibody (Ab) that binds to
CC one of the human proteins (P) designated PRO213, PRO1330, PRO1449,
CC PRO237, PRO324, PRO351, PRO362, PRO615, PRO531, PRO538, PRO3664, PRO618,
CC PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions
CC and methods for the diagnosis and treatment of neoplastic cell growth
CC and proliferation in mammals, including humans. Genes and polypeptides
CC encoded by them, that are amplified in the genome of a tumour cell, can
CC be identified and are useful targets for the treatment and prevention of
CC certain cancers and may be used to monitor tumour treatment. Compounds
CC that inhibit the expression or activity of the identified polypeptides
CC can be identified and used as antagonists. Benign or malignant tumours,
CC inflammatory disorders and immunological disorders can be treated.
CC AAC58123 to AAC58224 represent hybridisation probes and PCR primers used

CC in the isolation of the human PRO sequences. AAC58225 to AAC58241 and
CC AAB24041 to AAB24056 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.

XX
SQ Sequence 802 AA;

Alignment Scores: 3.64e-44 Length: 802
Pred. No.: 489.50 Matches: 99
Score: 489.50
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 21 Gaps: 5

US-09-735-713A-1 (1-921) x AAB24052 (1-802)

QY 160 AGTCCGATTCCTGGAGGAGCCAAAGTGGAGAGGGTTCCTATCTCCCTGGCAGGATATCTCTG 219
|||||
Db 566 SerArgIleValGlyGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGCGAGCAAGCATATTTGGAGGAGCATCGTCTCACACAGTGGGTGATCAGG 279
|||
Db 586 GlnValArgGlyArgHisIleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleHr 605
QY 280 GCGGCTCACTGCATTCGCAACAGAACATTTGTCTACTTTG-----AATGTTACTGCT 333
|||||
Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTAGCTTAAGCCAGACAGACCAGAGAGCAAACTCTCATTATTTAACTGTC 393
|||||
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATACATCCACATTTCTCCACCAGAACCAATGGACTATGATATGCCCTTTTGAAG 453
|||
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664
QY 454 ATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTTCCAGAGCTGCGG 513
|||
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GAGCAATTTGAGCGTGGTATTTATTTGTACAAGTGGAGCTGGCGCGCTTAAGTGAAGT 573
|||||
Db 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
QY 574 GCGCTCTCTCAAGCTCTTGAGGAGTGAATCTGCCCTATTTTGACCTGGGAAGAGTGT 633
|||
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
QY 634 GTGGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGGAGACCTTTCTTTTCACAGGT 693
|||
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCCTGATGGAGGAGAGACCATGTCAGGAGATTCAGGAGGTTCACTCATGTGCCGG 753
|||
Db 741 TyrArgLysGlyLysAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 813
|||
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GGCTGGAGAAACAATGTGAGGAAAAGTATGATCAAGGATCCCT-----GGGATCTTC 864
|||
Db 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATTAGTAAGTCTTCTCTGGATCCAGGAA 900
|||
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

Search completed: March 25, 2003, 06:58:03
Job time : 190 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 25, 2003, 07:00:26 ; Search time 51 Seconds

(without alignments)
1930.931 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 1668

Sequence: 1 atgagtctcaaatgcttat.....acatcccaactggtactaa 921

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Delop 6.0 , Delext 7.0

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 442306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USP70_spool/US09735713/runat_18032003_124356_23271/app_query.fasta_1.1095
-DB=PublishedApplications_AA -QFWT=fastan -SUFFIX=rapp -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	97.5	306	9	US-09-735-713A-2
2	1608	96.4	302	9	US-09-735-713A-4
3	1356	81.3	556	10	US-09-888-615-106
4	898	53.8	164	9	US-09-735-713A-6

5	489.5	29.3	802	9	US-09-978-295A-169	Sequence 169, App
6	489.5	29.3	802	9	US-09-978-697-169	Sequence 169, App
7	489.5	29.3	802	9	US-09-978-192A-169	Sequence 169, App
8	489.5	29.3	802	9	US-09-999-832A-169	Sequence 169, App
9	489.5	29.3	802	9	US-09-978-189-169	Sequence 169, App
10	489.5	29.3	802	9	US-09-978-608A-169	Sequence 169, App
11	489.5	29.3	802	9	US-09-978-191A-169	Sequence 169, App
12	489.5	29.3	802	9	US-09-978-403A-169	Sequence 169, App
13	489.5	29.3	802	9	US-09-978-564A-169	Sequence 169, App
14	489.5	29.3	802	9	US-09-978-585A-169	Sequence 169, App
15	489.5	29.3	802	9	US-10-017-081A-169	Sequence 169, App
16	489.5	29.3	802	10	US-09-888-615-113	Sequence 113, App
17	467	28.0	453	9	US-09-978-295A-69	Sequence 69, Appl
18	467	28.0	453	9	US-09-978-697-69	Sequence 69, Appl
19	467	28.0	453	9	US-09-978-192A-69	Sequence 69, Appl
20	467	28.0	453	9	US-09-999-832A-69	Sequence 69, Appl
21	467	28.0	453	9	US-09-978-189-69	Sequence 69, Appl
22	467	28.0	453	9	US-10-174-590-64	Sequence 64, Appl
23	467	28.0	453	9	US-10-176-758-64	Sequence 64, Appl
24	467	28.0	453	9	US-10-175-737-64	Sequence 64, Appl
25	467	28.0	453	9	US-10-173-706-64	Sequence 64, Appl
26	467	28.0	453	9	US-10-175-738-64	Sequence 64, Appl
27	467	28.0	453	9	US-10-175-752-64	Sequence 64, Appl
28	467	28.0	453	9	US-10-176-482-64	Sequence 64, Appl
29	467	28.0	453	9	US-10-176-757-64	Sequence 64, Appl
30	467	28.0	453	9	US-10-176-913-64	Sequence 64, Appl
31	467	28.0	453	9	US-10-180-552-64	Sequence 64, Appl
32	467	28.0	453	9	US-10-180-557-64	Sequence 64, Appl
33	467	28.0	453	9	US-10-173-700-64	Sequence 64, Appl
34	467	28.0	453	9	US-10-174-572-64	Sequence 64, Appl
35	467	28.0	453	9	US-10-174-579-64	Sequence 64, Appl
36	467	28.0	453	9	US-10-174-582-64	Sequence 64, Appl
37	467	28.0	453	9	US-10-174-588-64	Sequence 64, Appl
38	467	28.0	453	9	US-10-175-739-64	Sequence 64, Appl
39	467	28.0	453	9	US-10-175-740-64	Sequence 64, Appl
40	467	28.0	453	9	US-10-175-743-64	Sequence 64, Appl
41	467	28.0	453	9	US-10-176-488-64	Sequence 64, Appl
42	467	28.0	453	9	US-10-176-492-64	Sequence 64, Appl
43	467	28.0	453	9	US-10-176-747-64	Sequence 64, Appl
44	467	28.0	453	9	US-10-176-750-64	Sequence 64, Appl
45	467	28.0	453	9	US-10-176-985-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-09-735-713A-2
; Sequence 2, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abulin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713A-2

Alignment Scores:
Pred. No.: 1.91e-152 Length: 306

Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 97.48% Indels: 0
DB: 9 Gaps: 0

US-09-735-713A-1 (1-921) x US-09-735-713A-2 (1-306)

```
Qy 1 ATGAGTCTCAAAATGCTTATTAAGCAGGAAACAAGCTGATTTTACTAGGAATAGTCTTT 60
Dy 1 MetSerLeuLysMetLeuLeuLeuSerArgAsnLysLeuLeuLeuLeuLeuGlyValPhe 20
Qy 61 TTTGAACRAGGTAATCTGCACTCTTTTCGCTCCCAAGCTCCAGTGTGGCAGAGT 120
Dy 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Qy 121 CTGTTAAGGTACAGCTTGAATATTATTAACATTTTCAGTCCCATCTTTGGAGGAGC 180
Dy 41 LeuValLysValGlnProTrpAsnTyrPheAsnLeuPheSerArgLeuLeuGlyGlySer 60
Qy 181 CAAGTGGAGAAAGGTTCTATCCCTGCGAGGTATCTCTGAAACAAAGGCAAGCATATT 240
Dy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Qy 241 TGTGGAGAAAGCATCGTCTACACAGTGGGTGATCAGCGGGCTCACTGCATTGCCAAAC 300
Dy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Qy 301 AGAACATGTGCTACTTTTCAATGTTACTCTGGAGAGTATGACTTAAGCCACAGAC 360
Dy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Qy 361 CCAGGAGAGCAACTCTCACATATTGAACACTGTCATCATACATCCACATTTCTCCACCAAG 420
Dy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleLeuHisProHisPheSerThrLys 140
Qy 421 AAACCAATGGACTATGATATTCCTTTTGAAGTGGCTGGAGCTTCCAAATTTGGCCAC 480
Dy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Qy 481 TTTCTGGGGCCCATATCTTCAGAGCTCGGGAGCAATTTGAGGCTGGTATTATTGT 540
Dy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Qy 541 ACAACTCAGGCTGGGGCGGCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTCAGGAA 600
Dy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200
Qy 601 GTGAATCTGCCTATTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGG 660
Dy 201 ValAsnLeuProIleLeuThrTrpGluCysValAlaLeuLeuThrLeuLysArg 220
Qy 661 CCCATCAGTGGGAAGACTTTCTTTTGACAGCTTTTCTGATGAGGAGGAGAGCGCATGT 720
Dy 221 ProfilesGlyLysThrPheLeuCysThrGlyPheProaspGlyGlyArgAspAlaCys 240
Qy 721 CAGGAGATTCCAGGAGTTCACTCATGTGCCGGAATTAAGAAAGGGCGCTGGACTCTGGCT 780
Dy 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Qy 781 GGTGTGACTTCTCGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 840
Dy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
Qy 841 GATCAAGATCCCTCGGAATCTTCACAGACATTAAGTAAAGTCTTCTCGATCCAGCAAG 900
Dy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Qy 901 CACATCCAAACTGGTAAC 918
Dy 301 HisIleGlnThrGlyAsn 306
```

RESULT 2

US-09-735-713A-4
; Sequence 4, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Waite, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1 Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09735.713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713A-4

Alignment Scores:
Pred. No.: 1,15e-150 Length: 302
Score: 1608.00 Matches: 300
Percent Similarity: 99.34% Conservative: 0
Best Local Similarity: 99.34% Mismatches: 2
Query Match: 96.40% Indels: 0
DB: 9 Gaps: 0

US-09-735-713A-1 (1-921) x US-09-735-713A-4 (1-302)

```
Qy 13 ATGCTTATAAGCAGGAAACAAGCTGATTTTACTAGGAATAGTCTTTTGAACRAGGT 72
Dy 1 MetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhePheGluArgGly 20
Qy 73 AAATCTGCARCTCTTTCTGCTCCCAAGCTCCAGTGTGGCAGAGTCTGGTAAAGTA 132
Dy 21 LysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal 40
Qy 133 CAGCTTTGGAAATTTTAACTTTTTCAGTGCATCTTGGAGGAAAGCAAGTGGAGAG 192
Dy 41 GlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlySerGlnValGlnLys 60
Qy 193 GGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGAGAGCATATTTGTGGAGGAGC 252
Dy 61 GlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlySer 80
Qy 253 ATCGTCTCACCACAGTGGGTGATCAGCGCGCTCACTGCATTCGCAACAGAGAAATTTGTG 312
Dy 81 IleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal 100
Qy 313 TCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCACAGACAGAGAGAGCAA 372
Dy 101 SerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGln 120
Qy 373 ACTCTCACTATTTGAACACTGTCATCATCATCATCTCCACATTTCTCCACCAAGAAACAAATGGAC 432
Dy 121 ThrLeuThrIleGluThrValIleHisProHisPheSerThrLysLysProMetAsp 140
Qy 433 TATCATATTTGCCCTTTTGAAGATGGCTGGAGCTTCCAAATTTGCCACTTTGTGGGGCCC 492
Dy 141 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 160
Qy 493 ATATGCTCTCCAGAGCTGGGGAGCAATTTGAGGCTGGTATTATTGTACAACTGCAGGC 552
Dy 161 IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 180
Qy 553 TGGGGCCGCTTAAGTGAAGTGGGTGCTCTCTCAAGTCTTTGAGGAGGAGTGAATCTGCCT 612
Dy 553 TGGGGCCGCTTAAGTGAAGTGGGTGCTCTCTCTCAAGTCTTTGAGGAGGAGTGAATCTGCCT
```

```
Db 181 TrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnValAsnLeuPro 200
QY 613 ATTTGACCTGGAGAGTGTGGCAGCTCTGTTAACACTAAAGAGGCCCATCACTGGG 672
Db 201 IleLeuThrTrpGluGlyCysValAlaLeuLeuThrLeuLysArgProIleSerGly 220
QY 673 AAGACCTTTCTTTGACACAGGTTTCTGATGGAGGAGAGAGCATGTCCAGGAGATTCA 732
Db 221 LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSer 240
QY 733 GGAGTTCTACTCATGTGCCCGAATAAGAAAGGGCCCTGGACTCTGCTGGTGTGACTTCC 792
Db 241 GlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSer 260
QY 793 TGGGGTTTGGCTGTGGTGGAGCTGGAGAAACAATGTGAGGAAAGTGCATCAAGGATCC 852
Db 261 TrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGlnGlySer 280
QY 853 CCTGGGATCTTACACACATATTAGTAAAGTCTTCTCGGATCCACGAAACATCCAAACT 912
Db 281 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 300
QY 913 GGTAAC 918
Db 301 GlyAsn 302
RESULT 3
US-09-888-615-106
; Sequence 106, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY J.
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN.
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888, 615
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-106
```

```
Alignment Scores:
Pred. No.: 1,19e-125 Length: 556
Score: 1356.00 Matches: 262
Percent Similarity: 87.9% Conservative: 29
Best Local Similarity: 85.34% Mismatches: 8
Query Match: 81.29% Indels: 8
DB: 10 Gaps: 2
```

US-09-735-713A-1 (1-921) x US-09-888-615-106 (1-556)

```
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAACAAGCTGATTTTACTAGGAATAGTCTTT 60
Db 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20
QY 61 TTTGAAACRAGGTAATCTGCARCTCTTTCGCTCCCAAGAGTCCCAAGTGTGGGAGAGT 120
Db 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
QY 121 CTGGTTAAGGTACAGCCTTGGAAATATTATTTACATTTTTCAGTCGCATTTCTTGAGGAAGC 180
Db 41 LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer 60
```

```
QY 181 CAAGTGGAGAAGGTTTCCTTATCCCTGGCAGGTATCTCTGAAACAAAGGAGCAGCATATT 240
Db 61 GlnValGluGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
QY 241 TGTGGAGGAAGCATCGTCTCACACAGTGGGTGATCACGGCGGCTCACATTCGCAATTCGCAAC 300
Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
QY 301 AGAAACATTTGCTCTACTTTTGAATCTTACTGCTGGAGAGATATGACTTAAAGCCAGACAGAC 360
Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
QY 361 CCAGAGAGCAAACTCTCACTATTGAAACTGTCACTCATACATCCACATTTCTCCACCAAG 420
Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
QY 421 AAACCAATGGACTATGATATATTCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCAC 480
Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
QY 481 TTTGTGGGCCCATATGTCTTCCAGAGCTGCGGAGCAATTTGAGGCTGTGTTTATTTCGT 540
Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
QY 541 ACAACTGCAGGCTGGGGCCGCTTAAGTGGCGTCTCTCTCAAAAGTCTTTCAGAGAA 600
Db 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
QY 601 GTGAATCTCCCTATTTTGACCTGGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
Db 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
QY 661 CCCATCAGTGGGAAGACCTTTCTTCCACAGCTTTTCTGATGGAGGAGAGACGCATCT 720
Db 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
QY 721 CAGGAGATTCAGGAGTTCACATCATGTGCCGGAATAAGAAAGGCGCTCGGCT 780
Db 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpAsp---Ser 259
QY 781 GGTGTGACTTCTCCGTTGGGCTGTGTCGAGGCTGTGTCGAGGAGAAACANTGTGAGGAAAGT 840
Db 260 GlyTrpSerIleTrpGluAlaGlnValGlySerLeuGluSerArgSerArgPro 279
QY 841 GATCAAGGATCCCT-----GGGATCTTCCACAGACATTTAGTAAA 879
Db 280 SerLeuGlyAsnLysValArgLeuCysLeuThrAsnAsnPhePheLysLysLeuAlaGly 299
QY 880 GTGCTTTCTCGATCCACGAA 900
Db 300 CysGlyThrTrpCysSerGlu 306
RESULT 4
US-09-735-713A-6
; Sequence 6, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
```

```

; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713A-6

Alignment Scores:      1.53e-80      Length:      164
Pred. No.:      898.00      Matches:      164
Score:      100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      53.84%      Indels:      0
DB:      9      Gaps:      0

US-09-735-713A-1 (1-921) x US-09-735-713A-6 (1-164)
QY 427 ATGACATGATATGCCCCCTTTTCAAGATGGCTGGAGCCCTCCAAATTTGCCACCTTTGTG 486
Db 1 MetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 20
QY 487 GGGCCCATATCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTATTGTACAACT 546
Db 21 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 40
QY 547 GCAGGCTGGGGCCCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTCAGGAAAGTGAAT 606
Db 41 AlaGlyTrpGlyArgLeuThrGluGlyCysValLeuSerGlnValLeuGlnGluValAsn 60
QY 607 CTGCTATTATTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACATAAGAGGCCCATC 666
Db 61 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 80
QY 667 AGTGGGAAGACCTTCTTTGACAGGTTTCTCTGATGGAGGAGAGCATGTGCAGGA 726
Db 81 SerGlyLysThrPheLeuCysThrGlyPheProaspGlyArgaspAlaCysGlnGly 100
QY 727 GATTACAGAGGTTCACTCATGTGCCGAATAAGAAAGGGCCCTGGACTGTGGCTGGTGTG 786
Db 101 AspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyVal 120
QY 787 ACTTCCCTGGGTTGGCTGTGTCGAGGTCGGAGGAGAACATGTGAGAGAAAGTGATCAA 846
Db 121 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln 140
QY 847 GGATCCCTGGGATCTTCACAGACATAGTAAAGTGTTCCTGGATCCCGAACACATC 906
Db 141 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 160
QY 907 CAAACTGGTATAC 918
Db 161 GlnThrGlyAsn 164

RESULT 5
US-09-978-295A-169
; Sequence 169, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
```

;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.: 6.72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-295A-169 (1-802)

QY 160 AGTCGATTCTTGGAGAGCCCAAGTGGAGAGGGTTCTCTATCCCTGGCAGGTATCTCTG 219
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 566 SerArgIleValGlyGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAAGGCAGAGCATATTTGTGGAGGAAGCATGCTCCACACAGTGGGTGATCAG 279
::: ||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 586 GlnValArgGlyArgHisIleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGGCTCACTGCATTCGAAACACAAACATGTGCTCTACTTTG-----AATGTTACTGCT 333
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGGAGAGCAAACTCTACTATTGAACTGTC 393
||||| :||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATACATCCACATTTCTCCCAAGAACCAATGACTATGATATGCTCCCTTTTGAAG 453
::: :||||| :||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTrpAspValAlaLeuGln 664
QY 454 ATGGCTGGAGCCTTCCAAATTTGGCCCACTTTGTGGGCCCATATGTCTTCCAGAGTCGCG 513
::: :||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GACAAATTGAGCGCTGGTTTATTTTACAACTGCAGCGTGGCGCGCTTAACCTGAAGGT 573
||||| :||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyAlaLeuArgGluGly 704

;
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15

;
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 6,72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservatives: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-697-169 (1-802)

QY 160 AGTCGCATTCTTGGAGGAACCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTG 219
Db 566 SerArgIleValGlyGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGCAGAGCATATTTGTGGAGGAAGCATGCTCTCACCACAGTGGGTGATCACG 279
Db 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGGCTCACTGCATTGCAACACAAACATTTGCTCTACTTTG-----AATGTTACTGCT 333
Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGAGAGCAAACTCTCATTATTGAACGTGC 393
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATATCCACATTTCTCCACCAAGAACCAATGAGCATATGATATATTGCCCTTTGAAG 453
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664
QY 454 ATGGCTGGAGCCCTTCCAATTTGCCCACTTTGTGGGGCCCATATGCTCTCCAGAGTCGCG 513
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GAGCAATTTGAGGCTGGTTTTATTGTACAACCTGCAGGCTGGGGCCGCTTTAACTGAAGT 573
Db 685 HisPhePheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
QY 574 GCGCTCTCTCAACAGTCTTGCAGGAAGTGAATCTGCCTATTTTGACCTGGGAAGAGTGT 633
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
QY 634 GTGGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACACAGT 693
Db 725 SerGluAlaTrpArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCTCTGATGGAGGAGAGACGATGTCAGGAGATTTCAGGAGGTTCACCTCATGTGCCGG 753
Db 741 TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCTCGGGGTTGGGCTGTGGTCTGA 813
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GGCTGGGAACAATGTGAGGAAAAGTATCAAGGATCCCT-----GGGATCTTC 864
Db 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATTAGTAAAGTCTTTCTCTGATCCAGGAA 900
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 7
US-09-978-192a-169
; Sequence 169, Application US/09978192a

Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29

```

; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.:      6,72e-40      Length:      802
Score:          489.50      Matches:      99
Percent Similarity: 56.75%      Conservative: 44
Best Local Similarity: 39.23%      Mismatches:  86
Query Match:      29.33%      Indels:      23
DB:               9          Gaps:         5

US-09-735-713A-1 (1-921) x US-09-978-192A-169 (1-802)
QY 160 AGTCGATCTTGGAGGAGCAAGTGGAGAGGGTCTCTATCCCTGGGAGGTATCTGTG 219
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 566 SerArgileValglyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGCAAGCATATTTGTGGAGGACATCGTCTCACCAGTGGGTGATCAGC 279
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 586 GlnValArglyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 605
QY 280 GCGGCTACTGCTGCAATGCAACAGAAATTTGTCTACTTTG-----AATGTTACTGCT 333
```

```

Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGGAGACAACTCTACTATTGAACTGTC 393
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 394 ATCATACATCCACATTTCTCCCAAGAAACCAATGAGTATGATATATGCCCTTTTGAAG 453
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuGln 664
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 454 ATGGCTGGAGCCTTCCAATTTGGCCCACTTTGTGGGCCCATATGTCTTCCAGAGCTCGG 513
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 514 GAGCAATTTGAGCGTGGTATTATTGTACAACGTGAGGCTGGGGCCGCTTAAGTGAAGGT 573
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 574 GCGTCTCTCAACAAGTCTTGCAGGAAGTGAATCTGCCTATTATTGACCTGGGAAGAGTGT 633
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 634 GTGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTGCACAGGT 693
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 694 TTTCTGTATGGAGGGAGAGACCGATGTCAGGGAGATTCAGGAGGTTCATCATGTCCGG 753
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 741 TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 754 AATAAGAAAGGGCGCTGACTCTGGCTGGTGTGACACTTCTGGGTTTGGCTGTGTCGA 813
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 814 GCGTGAGAAACAATGTGAGGAAAGTATCAGGATCCCT-----GGATCTTC 864
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 781 -----ProAsnTyrPheGlyValTyr 787
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 865 ACAGACATTTAGTAAAGTGTCTTCTCGATCCAGTCCAGAA 900
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 788 ThrArgileThrGlyValIleSerTrpIleGln 799
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
US-09-999-832A-169
; Sequence 169, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
```

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C63
CURRENT APPLICATION NUMBER: US/09/999,832A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600

```

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085382
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085380
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.:      6,72e-40      Length:      802
Score:          489.50      Matches:      99
Percent Similarity: 56.75%      Conservative: 44
Best Local Similarity: 39.29%      Mismatches: 86
Query Match:      29.35%      Indels:      23
DB:              5          Gaps:      5

US-09-735-713a-1 (1-921) x US-09-999-832a-169 (1-802)

QY 160 AGTCGCATCTTGGAGAACCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTG 219
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 566 SerArgileValGlyAlaValSerSerGluGlyCluTrpProTrpGlnAlaSerLeu 585

QY 220 AACAAAGGAGAGCATATTTGTGGAGGAGCATGCTCTCACACAGTGGGTGATCAG 279
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 586 GlnValArgGlyHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605

QY 280 GCGGCTCATCGCATTCACACAGAACATGTCCTACTTTG-----AATGTTACTGCT 333
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625

QY 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATTGAAACTGTC 393
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645

QY 394 ATCATCATCCCATTTCTCCACACAGAACCAATGGACTATGATATTCCTGCTTTTGAAG 453
    :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 646 LeuLeuHisPro--TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664

QY 454 ATGCTGGACCCCTCCAATTTGCCACTTTGTGGGCCCCATATGCTCTCCAGAGCTCCGG 513
    :|||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684

QY 514 GAGCAATTTGAGGCTGCTTTTATTGTGTACAACTGCAGGCTGGGGCCGCTTAAGTGAAGGT 573
    |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704

QY 574 GCGCTCTCTCACAGTCTTTGAGGAGTGAATCTGCTATTTTGGCTGGGAGAGTGT 633
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724

QY 634 GTGCAGCTCTGTAACTAAAGAGGCCCATCAGTGGGAGACCTTTCTTTTCACAGGT 693
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740

QY 694 TTTCTCTGATGGAGGAGACGATGTGAGGAGATTCAGGAGTTCACGTGCTGCGG 753
    :|||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 741 TyrArgLysGlyLysAspAlaCysGlnGlyAspSerGlyProLeuValCysLys 760

QY 754 AATAAGAAAGGCGCTGGACTGGCTGGTGTGACTTCTCTGGGTTGGGCTGTGGTCTGA 813
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780

QY 814 GGCTGGAGAAACAATGTGAGGAAAGTGATCAAGGATCCCT-----GGATCTTC 864
    |||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 781 -----ProAsnTyrPheGlyValTyr 787

QY 865 ACAGACATTAGTAAGTCTTCTCTGGATCCAGAA 900
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 9
US-09-978-189-169
; Sequence 169, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C7
; CURRENT APPLICATION NUMBER: US/09/978,189
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
```

1	PRIOR APPLICATION NUMBER: 60/078004
2	PRIOR FILING DATE: 1998-03-13
3	PRIOR APPLICATION NUMBER: 60/078886
4	PRIOR FILING DATE: 1998-03-20
5	PRIOR APPLICATION NUMBER: 60/078936
6	PRIOR FILING DATE: 1998-03-20
7	PRIOR APPLICATION NUMBER: 60/078910
8	PRIOR FILING DATE: 1998-03-20
9	PRIOR APPLICATION NUMBER: 60/078939
10	PRIOR FILING DATE: 1998-03-20
11	PRIOR APPLICATION NUMBER: 60/079294
12	PRIOR FILING DATE: 1998-03-25
13	PRIOR APPLICATION NUMBER: 60/079656
14	PRIOR FILING DATE: 1998-03-26
15	PRIOR APPLICATION NUMBER: 60/079664
16	PRIOR FILING DATE: 1998-03-27
17	PRIOR APPLICATION NUMBER: 60/079689
18	PRIOR FILING DATE: 1998-03-27
19	PRIOR APPLICATION NUMBER: 60/079663
20	PRIOR FILING DATE: 1998-03-27
21	PRIOR APPLICATION NUMBER: 60/079728
22	PRIOR FILING DATE: 1998-03-27
23	PRIOR APPLICATION NUMBER: 60/079786
24	PRIOR FILING DATE: 1998-03-27
25	PRIOR APPLICATION NUMBER: 60/079920
26	PRIOR FILING DATE: 1998-03-30
27	PRIOR APPLICATION NUMBER: 60/079923
28	PRIOR FILING DATE: 1998-03-30
29	PRIOR APPLICATION NUMBER: 60/080105
30	PRIOR FILING DATE: 1998-03-31
31	PRIOR APPLICATION NUMBER: 60/080107
32	PRIOR FILING DATE: 1998-03-31
33	PRIOR APPLICATION NUMBER: 60/080165
34	PRIOR FILING DATE: 1998-03-31
35	PRIOR APPLICATION NUMBER: 60/080194
36	PRIOR FILING DATE: 1998-03-31
37	PRIOR APPLICATION NUMBER: 60/080327
38	PRIOR FILING DATE: 1998-04-01
39	PRIOR APPLICATION NUMBER: 60/080328
40	PRIOR FILING DATE: 1998-04-01
41	PRIOR APPLICATION NUMBER: 60/080333
42	PRIOR FILING DATE: 1998-04-01
43	PRIOR APPLICATION NUMBER: 60/080334
44	PRIOR FILING DATE: 1998-04-01
45	PRIOR APPLICATION NUMBER: 60/081070
46	PRIOR FILING DATE: 1998-04-08
47	PRIOR APPLICATION NUMBER: 60/081049
48	PRIOR FILING DATE: 1998-04-08
49	PRIOR APPLICATION NUMBER: 60/081071
50	PRIOR FILING DATE: 1998-04-08
51	PRIOR APPLICATION NUMBER: 60/081195
52	PRIOR FILING DATE: 1998-04-08
53	PRIOR APPLICATION NUMBER: 60/081203
54	PRIOR FILING DATE: 1998-04-09
55	PRIOR APPLICATION NUMBER: 60/081229
56	PRIOR FILING DATE: 1998-04-09
57	PRIOR APPLICATION NUMBER: 60/081595
58	PRIOR FILING DATE: 1998-04-15
59	PRIOR APPLICATION NUMBER: 60/081817
60	PRIOR FILING DATE: 1998-04-15
61	PRIOR APPLICATION NUMBER: 60/081819
62	PRIOR FILING DATE: 1998-04-15
63	PRIOR APPLICATION NUMBER: 60/081952
64	PRIOR FILING DATE: 1998-04-15
65	PRIOR APPLICATION NUMBER: 60/081838
66	PRIOR FILING DATE: 1998-04-15
67	PRIOR APPLICATION NUMBER: 60/082568
68	PRIOR FILING DATE: 1998-04-21
69	PRIOR APPLICATION NUMBER: 60/082569
70	PRIOR FILING DATE: 1998-04-21
71	PRIOR APPLICATION NUMBER: 60/082704
72	PRIOR FILING DATE: 1998-04-22
73	PRIOR APPLICATION NUMBER: 60/082804

1	PRIOR FILING DATE: 1998-04-22	
2	PRIOR APPLICATION NUMBER: 60/082700	
3	PRIOR FILING DATE: 1998-04-22	
4	PRIOR APPLICATION NUMBER: 60/082797	
5	PRIOR FILING DATE: 1998-04-22	
6	PRIOR APPLICATION NUMBER: 60/082796	
7	PRIOR FILING DATE: 1998-04-23	
8	PRIOR APPLICATION NUMBER: 60/083336	
9	PRIOR FILING DATE: 1998-04-27	
10	PRIOR APPLICATION NUMBER: 60/083322	
11	PRIOR FILING DATE: 1998-04-28	
12	PRIOR APPLICATION NUMBER: 60/083392	
13	PRIOR FILING DATE: 1998-04-29	
14	PRIOR APPLICATION NUMBER: 60/083495	
15	PRIOR FILING DATE: 1998-04-29	
16	PRIOR APPLICATION NUMBER: 60/083496	
17	PRIOR FILING DATE: 1998-04-29	
18	PRIOR APPLICATION NUMBER: 60/083499	
19	PRIOR FILING DATE: 1998-04-29	
20	PRIOR APPLICATION NUMBER: 60/083545	
21	PRIOR FILING DATE: 1998-04-29	
22	PRIOR APPLICATION NUMBER: 60/083554	
23	PRIOR FILING DATE: 1998-04-29	
24	PRIOR APPLICATION NUMBER: 60/083558	
25	PRIOR FILING DATE: 1998-04-29	
26	PRIOR APPLICATION NUMBER: 60/083559	
27	PRIOR FILING DATE: 1998-04-29	
28	PRIOR APPLICATION NUMBER: 60/083500	
29	PRIOR FILING DATE: 1998-04-29	
30	PRIOR APPLICATION NUMBER: 60/083742	
31	PRIOR FILING DATE: 1998-04-30	
32	PRIOR APPLICATION NUMBER: 60/084366	
33	PRIOR FILING DATE: 1998-05-05	
34	PRIOR APPLICATION NUMBER: 60/084414	
35	PRIOR FILING DATE: 1998-05-06	
36	PRIOR APPLICATION NUMBER: 60/084441	
37	PRIOR FILING DATE: 1998-05-06	
38	PRIOR APPLICATION NUMBER: 60/084637	
39	PRIOR FILING DATE: 1998-05-07	
40	PRIOR APPLICATION NUMBER: 60/084639	
41	PRIOR FILING DATE: 1998-05-07	
42	PRIOR APPLICATION NUMBER: 60/084640	
43	PRIOR FILING DATE: 1998-05-07	
44	PRIOR APPLICATION NUMBER: 60/084598	
45	PRIOR FILING DATE: 1998-05-07	
46	PRIOR APPLICATION NUMBER: 60/084600	
47	PRIOR FILING DATE: 1998-5-07	
48	PRIOR APPLICATION NUMBER: 60/084627	
49	PRIOR FILING DATE: 1998-05-07	
50	PRIOR APPLICATION NUMBER: 60/084643	
51	PRIOR FILING DATE: 1998-05-07	
52	PRIOR APPLICATION NUMBER: 60/085339	
53	PRIOR FILING DATE: 1998-05-13	
54	PRIOR APPLICATION NUMBER: 60/085338	
55	PRIOR FILING DATE: 1998-05-13	
56	PRIOR APPLICATION NUMBER: 60/085323	
57	PRIOR FILING DATE: 1998-05-13	
58	PRIOR APPLICATION NUMBER: 60/085582	
59	PRIOR FILING DATE: 1998-05-15	
60	PRIOR APPLICATION NUMBER: 60/085700	
61	PRIOR FILING DATE: 1998-05-15	
62	PRIOR APPLICATION NUMBER: 60/085689	
63	PRIOR FILING DATE: 1998-05-15	
64	PRIOR APPLICATION NUMBER: 60/085579	
65	PRIOR FILING DATE: 1998-05-15	
66	PRIOR APPLICATION NUMBER: 60/085580	
67	PRIOR FILING DATE: 1998-05-15	
68	PRIOR APPLICATION NUMBER: 60/085573	
69	PRIOR FILING DATE: 1998-05-15	
70	PRIOR APPLICATION NUMBER: 60/085704	
71	PRIOR FILING DATE: 1998-05-15	
72	PRIOR APPLICATION NUMBER: 60/085697	

;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580

;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 6, 72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713A-1 (1-921) x US-09-978-191A-169 (1-802)

QY 160 AGTCGCATCTTGGAGGAACCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTG 219
Db 566 SerArgIleValGlyGlyAlaValSerSerGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGCAGAGCATATTTGTGGAGGAAGCATCGTCTCACCACAGTGGGTGATCAGC 279
Db 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGGCTCACTGCATTCGAAACAGAACATTTGTGTACTTTG-----AATGTTACTGCT 333
Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGAGAGCAAACTCTCACTATTGAAACTGTC 393
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATACATCCACATTTCTCCACCAGAAACCAATGAGCATATGATATATGCCCCCTTTCAAG 453
Db 646 LeuLeuHisPro--TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664
QY 454 ATGGCTGGAGCCCTTCCAAATTTGGCCACTTTTGTGGGGCCCATATGCTCTCCAGAGCTCGG 513
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GAGCAATTTGAGGCTGGTTTTATTGTACAACTGCAGGCTGGGGCCCTTTAACTGAAGGT 573
Db 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuAlaGluGly 704
QY 574 GCGTCTCTCACAAAGTCTTTGAGGAAGTGAATCTGCCTATTATTGACCTGGGAAGAGTGT 633
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
QY 634 GTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTTCTTTCACAGGT 693
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCTCTGATGGAGGAGACGATGTCAGGAGATTTCAGGAGGTTCAGAGGTTCATGTGCCGG 753
Db 741 TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGGCCCTGGACTCTGGGTGGTGTGACTCTCCCTGGGGTTGGGCTGTGTCGA 813
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GGCTGGGAACAACATGTGAGGAAAGTATCAAGGATCCCT-----GGGATCTTC 864
Db 781 -----ProAsnTyrPheGlyValTyr 787

QY 865 ACAGACATTAGTAAAGTCTTTCTCTGGATCCACGAA 900
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 12
US-09-978-403A-169
; Sequence 169, Application US/09978403A
; Publication No. US20030050240A1

GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC17
; CURRENT APPLICATION NUMBER: US/09/978,403A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545

;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 6,72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-403A-169 (1-802)

QY 160 AGTCGATCTTGGAGGAGCAAGTGGAGAGGGTCTATCCCTGGCAGGATCTCTG 219
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 566 SerArgIleValGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGCAAGCATATTTGTGGAGCAAGCTGCTCACCACAGTGGGTGATCAG 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGGCTCACTGCATGTCGAACAAACATTTGTGTACTTTG-----AATGTTACTGCT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 606 AlaAlaHisCysPheGlnGluAspSerMetalaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGAGAGCAAACTCTCCTACTATTGAACTGTC 393
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATACATCCCATTTCTCCACCAAGAAACCAATGGAGCTATGATATGCTTCCCTTTTGAAG 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuGln 664
QY 454 ATGGCTGGAGCCTTCCAAATTTGGCCACTTTGGGGCCCATATGCTTCCAGAGTGGCG 513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GAGCAATTTGAGGCTGGTTTATTTTGTACAACTGAGGCTGGGGCCGCTTAAGTGT 573
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
QY 574 GCGCTCTCTCAAGTCTTGCAGGAAGTGAATCTGCCTTATTTTGCCTGGGAAGAGTGT 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
QY 634 GTGCAGCTCTGTTAACTAAAGAGGCCCATCTACATGGAAGACCTTCTTCTTCCACAGGT 693
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCTGTATGGAGGAGGAGACCATGTGAGGAGATTCAGGAGTTTCACTATGTGCCGG 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 741 TyrArgLysGlyLysAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGGCGCTGCTCTGCTGTGTGACTTCTCTGGGTTCGGCTGTGTGCTCA 813
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GCGTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCCCT-----GGGATCTTC 864
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATTAGTAAGTGTCTTCTTCCTGGATCCACGAA 900
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 13

US-09-978-564A-169

; Sequence 169, Application US/09978564A

; Publication No. US20030050241A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C25
CURRENT APPLICATION NUMBER: US/09/978,564A
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918595
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 6, 72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservativeness: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-564A-169 (1-802)

QY 160 AGTCGATCTTGGAGAACCAAGTGGAGAGGGTCTCTATCCCTGGCAGGTATCTCTG 219
DB 566 SerArgileValGlyAlaValSerSerGlyGluTrpProtrpGlnAlaSerLeu 585
QY 220 AACAAGGACAGCATATTTGTGAGGAGCATCTCTCACCAGCTGGGTGATCAG 279
DB 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGGCTCACTGCTTCCAAACAGCATTTGCTCTACTTTG-----AATGTTACTGCT 333
DB 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGGAGCAAACTCTCACTATTGAACTGTC 393
DB 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATATCCACATTTCTCCACCAGAACCAATGAGTATGATATTCCTTTTGAAG 453
DB 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuGln 664
QY 454 ATGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGAGCTCGG 513
DB 665 LeuAspHisProValValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GACAAATTTGAGCTGGTTTATTTTGTACAACTGCAGGCTGGGGCCGCTTAACCTGAAGT 573
DB 685 HisPhePheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
QY 574 GCGCTCTCACAAGTCTTTCAGGAGGATGATCTGCTCTATTTGACCTGGGAAGAGTGT 633
DB 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
QY 634 GTGGCAGCTCTGTTAACTAAGAGGCCCATCATCAGTGGGAGACCTTTCTTTCACAGGT 693
DB 725 SerGluAlaTrpArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCCTGATGGAGGAGACGCATCTCAGGGAGATTCAGGAGGTTCACCTCATGTGCCGG 753

DB 741 TyrArgLysGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGGCGCTGGACTCTGGCTGTGTGACTTCTCTGGGTTGGGCTGTGTGCTCA 813
DB 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GCGTGGAGAAACAATGTGAGGAAAGTATCAAGGATCCCT-----GGGATCTTC 864
DB 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATTAGTAAAGTCTTCTCTGGATCCACGAA 900
DB 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 14

US-09-978-585A-169
; Sequence 169, Application US/09978585A
; Publication No. US20030049633A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC15
; CURRENT APPLICATION NUMBER: US/09/978,585A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 169
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-585A-169

Alignment Scores:
Pred. No.: 6, 72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservativeness: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-585A-169 (1-802)

QY 160 AGTCGATCTTGGAGAACCAAGTGGAGAGGGTCTCTATCCCTGGCAGGTATCTCTG 219
DB 566 SerArgileValGlyAlaValSerSerGlyGluTrpProtrpGlnAlaSerLeu 585

Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GGCTGGAGAAACAATGTGAGGAAAGTCAAGGATCCCT-----GGGATCTTC 864
Db 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATTAGTAAGTGCTTCTCGGATCCACGAA 900
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

Search completed: March 25, 2003, 07:24:21
Job time : 56 secs

THIS PAGE BLANK (USPTO)